FIRST EVIDENCE FOR VECTOR BOSON FUSION PRODUCED HIGGS BOSON IN THE WW^* DECAY CHANNEL AND SEARCH FOR EXTENDED HIGGS SECTOR

BY

PHILIP CHANG

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Physics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2017

Urbana, Illinois

Doctoral Committee:

Professor Kevin Pitts, Chair
Associate Professor Mark Neubauer, Director of Research
Professor Scott Willenbrock
Professor Brian Fields

Abstract

The discovery of a new boson state in 2012 shook the field of particle physics. The subsequent measurements of the properties of the newly found particle led to the conclusion that the particle is consistent the Standard Model Higgs boson within the uncertainty of the measurements. Individual observations of the Higgs boson production and decay channels are crucial in further confirming beyond a doubt that this is indeed the Standard Model Higgs boson. This thesis presents the first evidence for the vector boson fusion (VBF) produced Higgs boson in exclusive Higgs boson decay channel ($H \to WW^*$). The best single channel constraint on the fermionic coupling of the Higgs boson comes from the measurement presented in the thesis. The analysis is performed using an integrated luminosity of 25 fb^{-1} from $\sqrt{s} = 7$ and 8 TeV pp collision data recorded by the ATLAS detector at the LHC. The observed (expected) significance for the vector boson fusion produced $H \to WW^*$ is 3.2 (2.7) standard deviations. As the existence of the Higgs sector is established, using the Higgs boson as a tool, a strategy for a non-Standard Model decay of the Higgs boson $H \to ss \to bb\mu\mu$ is documented in this thesis, where s is a hypothetical (pseudo)scalar with connections to Dark Matter models. The strategy has been tested using data with an integrated luminosity of 36 fb^{-1} from $\sqrt{s} = 13$ TeV pp collisions recorded by the ATLAS detector and the sensitivity of the channel is analyzed. The expected upper limit on the branching ratio is found to be $\mathcal{B}(H \to bb\mu\mu) \leq 2.4 - 4.8 \times 10^{-4}$ at 95% confidence limit across the mass range of 20 GeV $< m_s <$ 60 GeV.

To my love So Young, Han, and Ethan

Acknowledgments

First and foremost, I thank my parents and family for providing me with an environment to think freely and appreciate the work of our Father. I am lucky to have such good family supporting me through the endeavor.

Beginning of my graduate school journey was tough. Despite living in Urbana-Champaign for many years prior to the graduate school and considering the city my home, I had difficulty finding my *scholarly home* where I can study and flourish freely until I met Mark.

I cannot thank enough to my advisor Mark Neubauer for helping me find my own research where I can strive. ATLAS collaboration is big and often intimidating for a beginning graduate student, yet Mark's continuous support throughout the years has grown me into an independent member of the collaboration. Mark has taught me the most invaluable skill of following one's *physics nose*.

I thank my colleagues in ATLAS collaboration who have helped me through the big collaboration. I reserve special thanks to Tae Min Hong and Ben Cerio for the awesome Triumvirate and also for all my colleagues from HSG3 VBF team. I also thank Markus Atkinson, Jian Cong, Viviana Cavaliere, and Lulu Liu for being awesome officemates.

But above all, I thank my wife for the unending support that knows no limit and for my lovely two sons for being just they way they are. My wife and my sons have traveled over the Atlantic ocean several times for my study and have sacrificed so much living in foreign places with a limited support. Through all the hardships and difficult times that I was going through the years, you have always shown me nothing but love and care. There is no one I can thank more than you, So Young. I love you always.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

2HDM	Two Higgs Double Model
ATLAS	A Toroidal LHC Apparatus
BSM	Beyond the Standard Model
BW	Breit-Wigner
BDT	Boosted Decision Tree
CERN	The European Organization for Nuclear Research
CL	Confidence Level
CMS	Compact Muon Solenoid
CR	Control Region
DM	Dark Matter
EF	Event Filter
EM	Electronmagnetic
EW	Electroweak
EWPO	Electroweak Precision Observable
EWSB	Electroweak Symmetry Breaking
FTK	Fast TracKer
ggF	gluon-gluon Fusion
GSF	Gaussian sum filter
HLT	High-level Trigger
IBL	Insertable B-Layer
ID	Inner Detector
JER	Jet Energy Resolution
JES	Jet Energy Scale
JVF	Jet Vertex Fraction
KLFit	Kinematic Likelihood Fit

LHC	Large Hadron Collider
LINAC	Linear Accelerator
LO	Leading Order
PDF	Parton Distribution Function
PID	Particle Identification
PS	Parton Shower
PS	Parton Shower
QCD	Quantum Chromodynamics
ROC	Receiver Operating Characteristics
RPC	Resistive Plate Chamber
UE	Underlying Event
SB	Side-band
SCT	SemiConductor Tracker
SR	Signal Region
SM	Standard Model
SSB	Second Stage Board
SSF	Second Stage Fit
TF	Transfer Function
TGC	Thin Gap Chamber
TRT	Transition Radiation Tracker
VBF	Vector Boson Fusion
VBS	Vector Boson Scattering

List of Symbols

e	Electron
μ	Muon
γ	Photon
κ	Higgs coupling modifier
κ_F	Fermionic Higgs coupling modifier
κ_V	Bosonic Higgs coupling modifier
η	Pseudorapidity
ϕ	Azimuthal angle
ΔR	Spatial separation in $\eta - \phi$ space
p_{T}	Transverse momentum
m_{T}	Transverse mass
O_{BDT}	Boosted Decision Tree algorithm output score
$E_{\mathrm{T}}^{\mathrm{miss}}$	Missing transverse energy
$p_{\mathrm{T}}^{\mathrm{miss}}$	Missing transverse momentum
ep	Electron-proton
pp	Proton-proton
$p\bar{p}$	Proton-antiproton
Γ	Particle width
$\mathcal{M}$	Scattering matrix
ℓ	Lepton (e or μ

Chapter 1

Introduction to Collider Physics

In the late 19^{th} and early 20^{th} century, particle physics as an endeavor flourished. Several elementary particles known today were discovered as rays of matter. The electron was first discovered in 1897 through the famous cathode ray experiment. Discoveries of alpha, beta, and gamma rays soon followed. These were the first source of accelerated particle beams of the time, before the advent of particle accelerators.

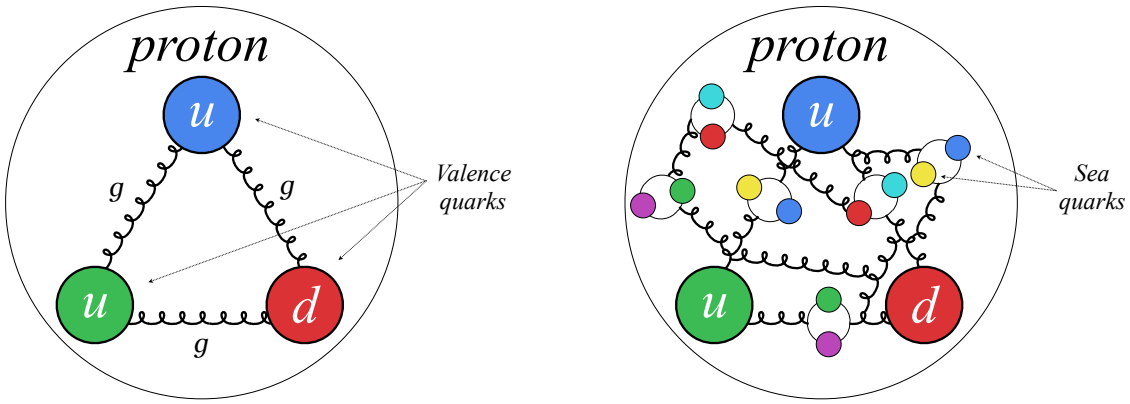
Rutherford and his team directed alpha particles at a thin gold foil and deduced the structure of the atom that it mostly consists of empty space with its nucleus in the center. Continuing his work, Rutherford shot alpha particles at various gasses and observed that hydrogen nuclei were contained in different gasses; Rutherford discovered that the hydrogen nucleus is a building block of all other atoms. Rutherford named this building block the *proton* in 1920 from the greek word *protos*, which means "first".

These hallmark experiments still illustrate the process of the present-day collider physics experiments, in which particles are collided with one another and through measurements of the scattering process one deduces the nature of the building blocks of the universe. Today, the role of the alpha rays illustrated in this example are replaced by various high energy particle beams produced by powerful accelerators around the world.

1.1 Protons

Protons are now understood as a composite particle built by *quarks* and held together by the strong interaction. There are six types of quarks: up, down, charm, strange, bottom and top. The strong interaction between quarks is mediated by the gluons. The main constituents of a proton are two up quarks and one down quark held together by exchanging gluons. The three quarks are called the *valence quarks*. This is illustrated by Fig. 1.1(a).

The exchanged gluons, through quantum fluctuations, may split into a quark anti-quark pair for a short period of time. As a result of this, a single proton can contain more than the three valence quarks at any given time. The creation and annihilation of quark anti-quark pairs through the gluon splitting are



(a) (b)

Figure 1.1: Illustration of proton structure. (a) shows only the valence quarks and the gluons being exchanged. In addition, (b) shows the sea quarks formed from vacuum polarization. At any given point in time, the proton will contain more quarks than just the three valence quarks.

illustrated in Fig. 1.1(b). Quarks created from the constant quantum fluctuation in the vacuum is referred as the *sea* quarks. The sea quarks can be any one of the six types of quarks; protons thus contain all types of quarks and gluons.

1.2 Standard Model

Everyday matter around us can be built by just a handful of elementary particles: up quark, down quark, and electron held together by gluons and photons. However, there are more elementary particles, in which together form the Standard Model (SM). Fig. 1.2 summarizes the particles in the SM. There are six quarks and six leptons, which are fermions, and each has a corresponding anti-particle of opposite charge. There force-carrying bosons, which mediates forces between the fermions. The quarks carry the color charges which allow quarks to interact with gluons. The quarks plus the charged leptons (i.e. electrons, muons, and taus) carry electroweak charges which allow them to interact with photons, $W^{\pm}$ and Z bosons. The neutrinos which are neutral only interact with the massive force carrier $W^{\pm}$ and Z bosons. The SM has provided an accurate prediction of the subatomic particles behavior for decades. As part of its many triumphs, the SM correctly predicted the existence of the Higgs boson which has been sought after for many years without success. In the year 2012, the last missing particle the Higgs boson has been finally discovered and the SM is now complete. The details of the Higgs boson physics in the context of the SM will be discussed in Chapter 2.

Despite its success, the SM is not without shortcomings. The SM fails to include gravitation and does

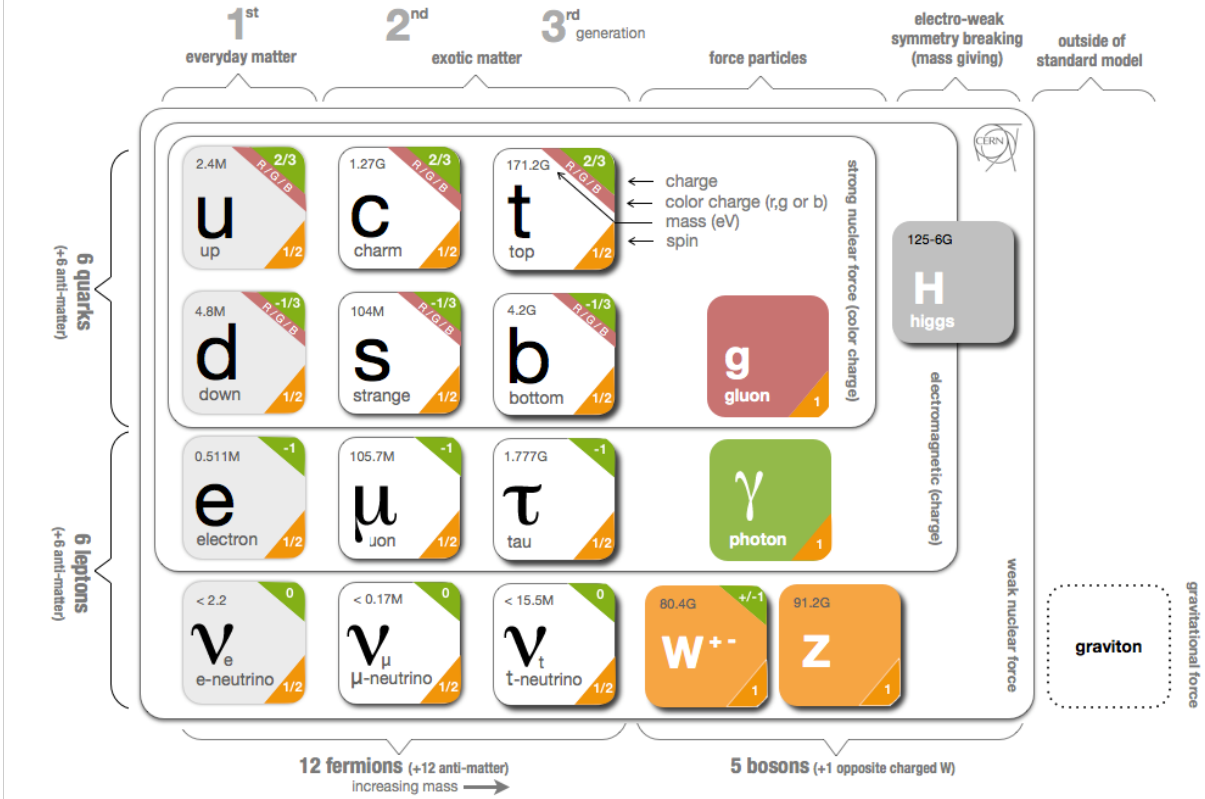


Figure 1.2: Summary of the Standard Model particles are shown. There are six quarks and six leptons, which together are the 12 fermions. The four force carrying bosons, g, γ, $W^{\pm}/Z$ mediates strong, electromagnetic, and weak interactions, respectively. Particles that interact with the three forces are grouped together by the round-edge rectangles. The fermions are labeled based on the charges they carry; the strong color charge is labeled with red color; the electroweak charge is labeled with green color. The spin is labeled by the orange color. The fifth boson in the Standard Model the Higgs boson is responsible for giving masses to the other particle. The last fundamental force gravity is not included in the SM (Image courtesy of CERN).

not provide a candidate for the Dark Matter (DM). The neutrinos are massless in the model, while several experiments have shown that neutrinos oscillate indicating neutrinos are not massless [1]. The SM also does not explain the observed asymmetry between the matter and anti-matter in the universe. There are also theoretical inconsistencies in the model that hint for unknown physics beyond the SM. The unsolved problems of the particle physics are being pursued through various experiments.

1.3 Large Hadron Collider

Hadron collider experiments use particle accelerators to accelerate hadrons, which consist of quarks, to a high energy and study the interaction between particles by colliding them. Large Hadron Collider (LHC) experiments are such examples with the world's currently largest particle accelerator the LHC. The LHC collides two separate proton beams with an unprecedented energy. By studying the debris from the proton scattering precisely, the SM is put to stringent tests.

The LHC is housed in a circular tunnel that is about 17 miles and about 330 ft underground. It is designed to accelerate protons up to the energy of 7 TeV, which corresponds to 14 TeV center-of-mass collision energy. Fig. 1.3 shows the aerial view of the LHC with relevant accelerator rings. After being stripped of its electrons, the protons are accelerated to 50 MeV through LINAC2. The beam is moved in steps from a smaller to a bigger synchrotron, accelerating protons to a higher energy through each step. At the Super Proton Synchrotron, the last step of acceleration before the beam is moved to the main LHC ring, the protons are accelerated to 450 GeV. Once injected to the main LHC ring, the protons are finally boosted to the collision energy through a series of radiofrequency cavities and steering magnetic dipoles. The accelerated protons are spaced out in bunches with each bunch containing about 100 billion protons. There are about 2200 bunches with a time interval of 25 ns in between.

There are four collision points along the main LHC ring. Each experiment's detector is housed at each collision points. Two insertion magnets composed of three quadrupoles surrounds each collision point focusing the proton beam into a tiny collision point; the magnet shrinks the beam of width 0.2 mm down to 16 μm for the collision. The focusing allows for a more dense beam, which leads to a larger number of pp collisions per bunch crossing. The maximum number of pp collisions recorded for a single bunch crossing by the ATLAS detector is 51 [2]. Each pp interaction point is called *primary vertex*. For most collision events, at most one primary vertex is of interest. Other non-interesting primary vertices are called *pile-up vertices* and the associated pp interactions are called *pile-up interactions*. The pile-up interactions introduce noisy background environment in the detector, which can affect the measurement of the interested primary vertex interaction.

Most pp collisions are glancing blows; the trajectory deviates only slightly from each proton's original path. When a "head-on" collision occurs, a burst of new particles are produced. Produced particles interact with the detector generating signals. The detector signals are recorded and the data is analyzed. An example event analyzed is shown in Fig. 1.4. This event is analyzed and has been concluded that one of the collision produced a Higgs boson decaying to two electrons (green tracks) and two muons (red tracks) through a pair of Z bosons. In addition to the Higgs boson, one can see a dense environment of tracks. One of the challenge

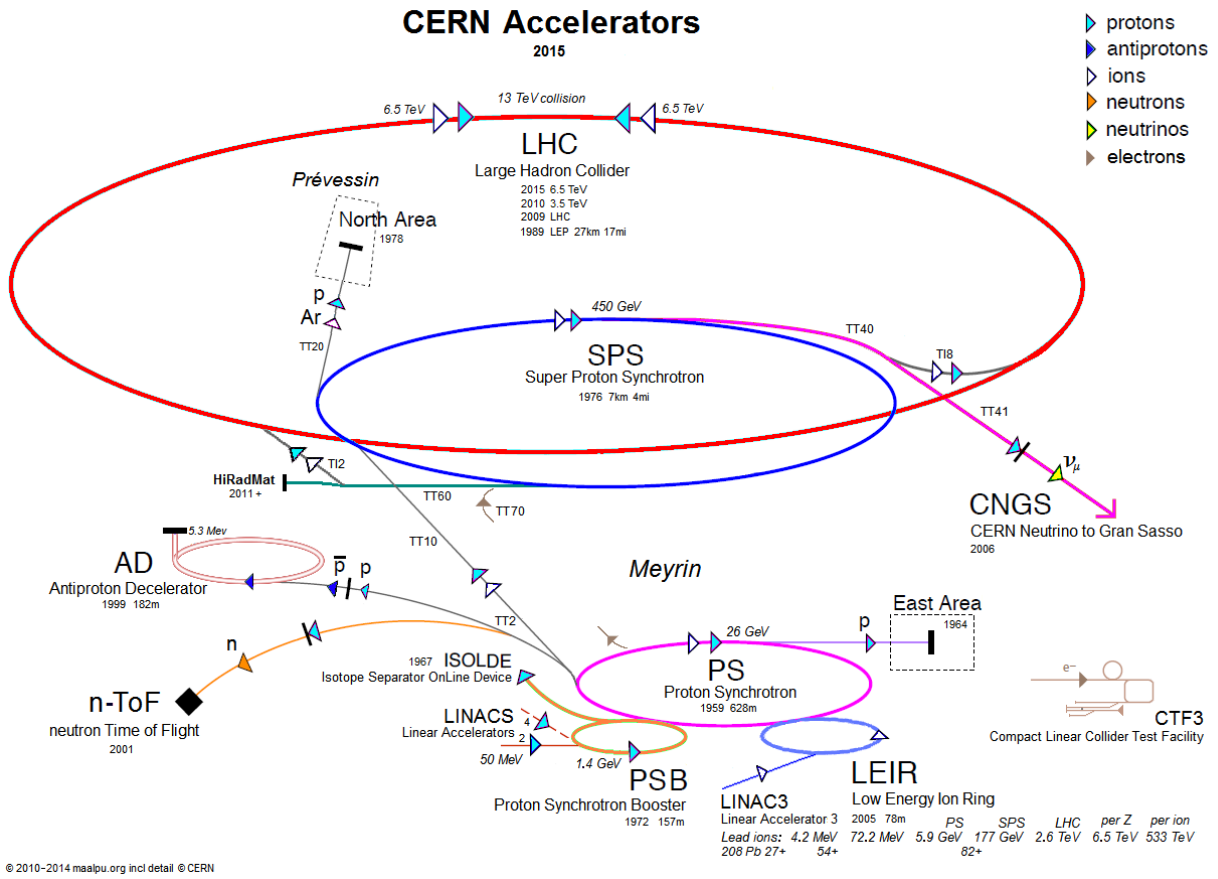


Figure 1.3: Summary diagram of the accelerator complexes at CERN. The Large Hadron Collider is shown as the largest circle. The operating energy of the particle beam for each accelerator is labeled along the line or the circumference (Image courtesy of CERN).

of the LHC experiment is analyzing events with large pile-up interactions, such as the one shown in Fig. 1.4.

In 2011 and 2012, the proton beams were accelerated to have an energy of 3.5 or 4 TeV; for 2015 and 2016 it was further raised to 6.5 TeV. The corresponding center of mass energies of the collisions are 7, 8, and 13 TeV, respectively. The 2011 and 2012 data taking is referred to as Run 1, and 2015 and 2016 data taking is referred to as Run 2. The Run 2 is planned to continue until the end of 2018.

1.4 Parton Distribution Function

For most head-on pp collisions, only one of the constituent from each proton interacts with one another. The colliding constituent can be a valence quark, a sea quark, or a gluon. A detailed description of the inner constituents, named *partons*, is essential in predicting the outcome of the pp collisions. To probe the structure of the protons, one performs a similar procedure as Rutherford and his team did, which is to direct particle beams at protons. Through the scattering experiments, the proton structure is probed. One of the

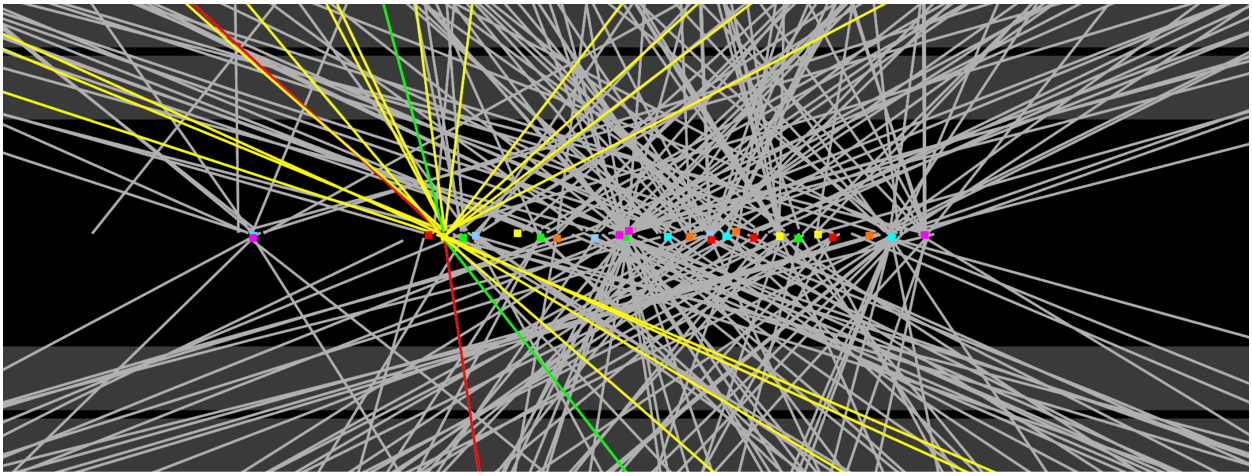


Figure 1.4: $H \to ZZ \to ee\mu\mu$ candidate event with 26 pp collisions. The green lines indicate the reconstructed track of the electrons from the Higgs boson decay. The red lines indicate the reconstructed track of the electrons from the Higgs boson decay. Other yellow lines are remaining charged particle tracks from the vertex associated with the Higgs boson decay. Gray tracks are other charged particles not associated with the Higgs boson decay vertex mainly from the other 25 pile-up pp collisions (ATLAS Experiment © 2017 CERN).

structural information obtained is called the parton distribution function (PDF).

The parton distribution function $f_i(x, Q^2)$ signifies the density of parton type i carrying longitudinal momentum fraction of x for a scattering experiment with momentum transfer Q. The definition of $f_i(x, Q^2)$ implies that,

$$\int_0^1 x f_i(x, Q^2) dx = (\text{fraction of momentum carried by parton type } i). \tag{1.1}$$

If summed over all parton types, one expects unity,

$$\sum_i^{\text{parton types}} \int_0^1 x f_i(x, Q^2) dx = 1. \tag{1.2}$$

Determining $f_i(x, Q^2)$ for all Q^2 values is a difficult task. The dynamics of the quarks and gluons are described by the quantum chromodynamics (QCD) and the theory becomes non-perturbative for low-energy. This leads to difficulties in predicting accurately the PDFs. The PDFs are thus obtained *a posteriori* through a global fit to various experimental data from ep, pp, and $p\bar{p}$ collisions probing various Q^2 regimes. The PDFs in the Q^2 regime not probed directly by the experiments are obtained through extrapolation from the Q^2 regime accessible by data. The evolution of PDFs from one Q^2 regime to another is described by DGLAP equation [3].

Fig. 1.5 shows the distribution of $x f_i(x, Q^2)$ for two different Q^2 values. The left plot with $Q^2 = 10$ GeV^2 shows the two distributions u_v and d_v peaking at high x values, where u_v and d_v are the PDFs for the valence

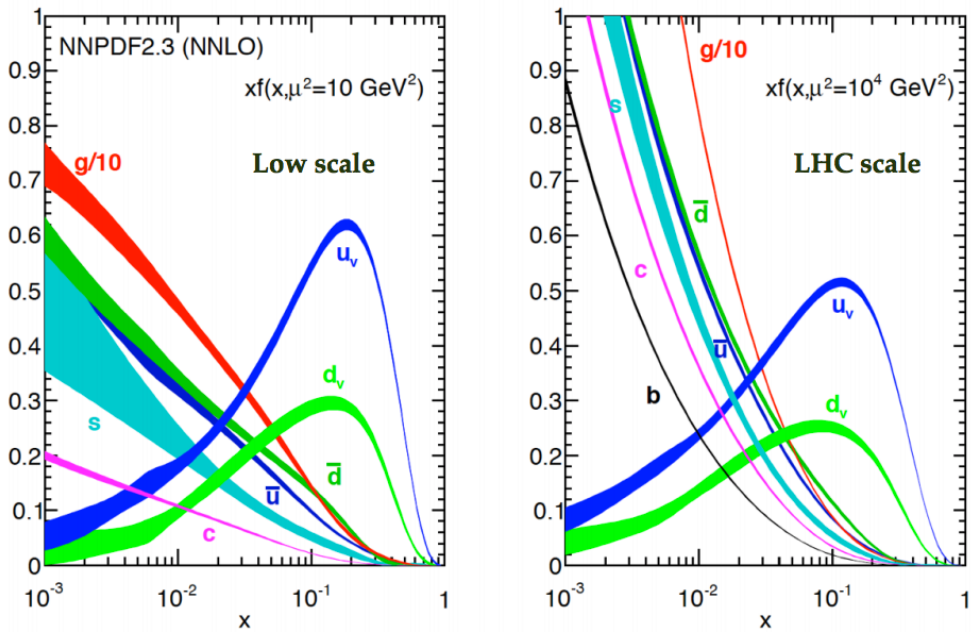


Figure 1.5: Parton Distribution Function (PDF) from Fig. 19.5 of Ref. [4]. (left) PDF for momentum transfer scale of $\mu^2 = 10$ GeV^2. Significant valence quark contributions are shown over the vast sea quark contributions. (right) PDF for momentum transfer scale of $\mu^2 = 10^4$ GeV^2, which is the scale more relevant at the LHC experiments. The sea quark contributions start to dominate over the valence quark contributions. The gluon contributions in both cases are very high as the plot shows the PDF for gluon scaled down by factor of 10. This implies that the proton collisions are dominated by gluon interactions.

up and down quark, respectively. The u_v distribution peaks twice as large as the d_v as there are two up quarks in a proton. The PDFs for the sea quarks (lines denoted as one of the quark flavor but without a subscript v in Fig. 1.5) show a falling distribution as x approaches unity. The PDFs for gluons are shown in red curve with the label $g/10$ and the "/10" signifies that the curve is scaled down by factor of 10. The right plot in Fig. 1.5 with $Q^2 = 10^4$ GeV^2 shows the PDF distributions more appropriate for the LHC experiments. The sea quarks and gluons contributions are relatively larger than at the low scale $Q^2 = 10$ GeV^2 yet the valence quark contributions are still dominant. As the dominant contribution from the pp collision is expected to be from the gluons, the LHC can be thought of as a *gluon collider*.

1.5 Basic Kinematics of Hadron Collider Physics

1.5.1 Coordinate System

All detectors at the LHC, with the exception of one, are cylindrical along the beampipe. The symmetry of the direction of beams dictates that the collisions are always symmetric in the azimuthal direction. The suitable coordinate system is thus a cylindrical one. Fig. 1.6 shows the coordinate system used in this thesis.

The origin is defined to be where collisions occur. The x-axis points towards the center of the LHC ring. The y-axis points towards the surface. The z-axis is pointing along the beam pipe, counter-clockwise when viewd from the top of the ring. The azimuthal angle ϕ is defined to be on the xy plane, which is defined to be the "transverse plane". The polar angle θ is defined following the usual convention of a cylindrical coordinate system.

1.5.2 Transverse Momentum Conservation

The partons inside each proton that collide carry off some unknown fractions of protons' momentum based on the PDF. The collision point is thus boosted along the z-axis by some unknown amount relative to the lab frame. The transverse momentum of the incoming protons, on the other hand, are negligibly small compared to the detector resolution and can be considered zero. The momentum conservation thus says that the vector sum of all outgoing particles momentum should equal to zero only along the transverse plane direction but not in the longitudinal direction. As a consequence, the transverse momentum p_{T} of a particle is an important quantity of the particle.

1.5.3 Rapidity and Pseudorapidity

The polar angle θ between two particles are different for two different inertial frames. As the boost along the z-axis is not known for each collision, the polar angle θ is not a useful quantity for the hadron collider physics. A separate quantity called the rapidity y is used instead,The rapiditiy is defined as,

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z} \tag{1.3}$$

where E is the energy of the particle and p_z is the momentum component along the z-axis. Another quantity often used is called the pseudorapidity and is defined as,

$$\eta = -\ln \tan \frac{\theta}{2}, \tag{1.4}$$

which is a more intuitive quantity than rapidity as it is monotonically transformed from the familiar polar coordinate θ. For massless particles the pseudorapidity is identical to the rapidity and therefore can be used interchangeably. For massive particles, the rapidity should be used instead as $E \neq |p|$. The geometrical distance in the $\eta - \phi$ space, $\Delta R \equiv \sqrt{\Delta\phi^2 + \Delta\eta^2}$, is often used to describe the degree of separation between two objects.

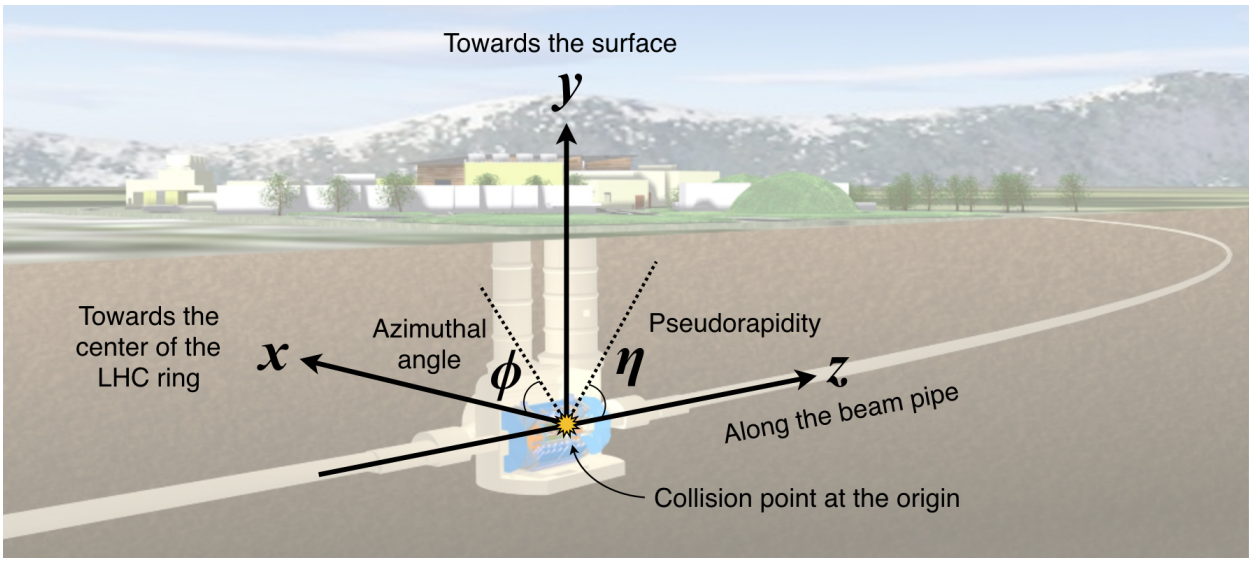


Figure 1.6: The coordinate system used commonly in ATLAS and throughout the thesis is illustrated. The x-axis is pointing towards the center of the LHC ring. The y-axis is pointing towards the surface. The z-axis is pointing along the beam pipe, counter-clockwise when viewd from the top of the ring. The azimuthal angle ϕ is defined to be the angle in the xy plane. The pseudorapidity η is a transformation from the polar angle θ. See texts for detail on the definition and discussion of η.

1.6 Rate of Physics Processes

The possible outcomes of pp collisions are near endless and not all outcomes are of interest. Most analyses search for rare signal events over large background processes, so rates of various processes must be understood. The quantity called cross section $\sigma_{pp\to X}$, where X denotes some final state configuration, is used as a measure of the rate of a given process and has a unit of an area*. The larger the cross section of a process, the more likely an event of the process occurs in a pp collision. Fig. 1.7 shows cross sections of important SM processes. The left most column shows that the total pp cross section σ_{total} is $\mathcal{O}(10^{11}$ pb). The second column shows that the W boson cross section is $\mathcal{O}(10^{5}$ pb), which is one of the largest cross section among the interesting SM processes, yet only takes up one part per million of the total pp cross section. It is also worth noting that the total Higgs boson cross section is one part per *billion* of the total pp cross section.

The number of expected events of a specific process is expressed in terms of the cross section as,

$$N_{pp\to X} = \sigma_{pp\to X} \times L, \tag{1.5}$$

where $N_{pp\to X}$ is the expected number of events for the process $pp \to X$, and L is the integrated luminosity.

*The reason for representing it as an area comes from the physics of classical scattering where the *cross section* of a target represents how likely the scattering will occur.

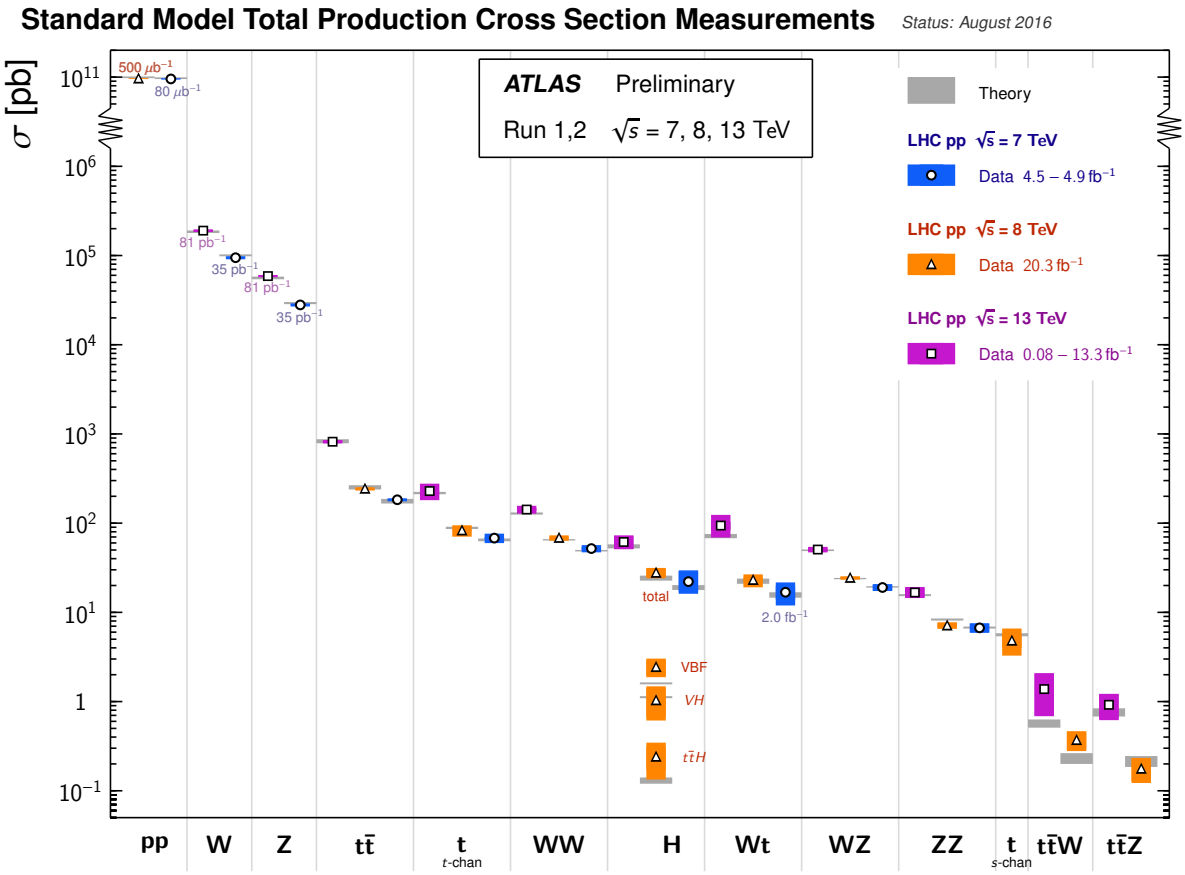


Figure 1.7: Cross sections of various Standard Model Processes. The y-axis shows the cross section of the process shown on each column in the unit of pb. The total pp collision cross section is shown to be around $\mathcal{O}(10^{11})$ pb. The first process with a relatively high-p_{T} lepton in the final state is the W boson production process. The W boson production cross section is around 2×10^5 pb, which is about 6 orders of magnitude smaller than the total pp collision cross-section. The total vector boson fusion (VBF) Higgs boson production cross section is around 2 pb, which is about 10 orders of magnitude smaller than the total pp collision cross-section [5].

The integrated luminosity L is,

$$L = \frac{N_{\mathrm{tot}}}{\sigma_{\mathrm{total}}}, \tag{1.6}$$

where N_{tot} is the number of all pp collisions occured during the data taking. The integrated luminosity is measured through several methods involving dedicated "luminometer" detectors, which monitors the on-going collisions.

1.7 Decay Branching Ratio

Most particles in the SM are unstable. Once an unstable particle is produced, it quickly decays to another particles. What is detected are only the relatively stable particles that traverse enough distances to interact and leave traces in the detector apparatus. For most decays in the SM, the decays are two-body decays and the rates for each different types of decays are different. The decay rate must be factored in when making a prediciton of how many events are expected in the recorded data.

The two-body decay of a parent particle with mass m_X to daughter particles with masses m_Y and m_Z with the coupling constant of g_c is shown in Fig. 1.8. The rate of the two-body decay is [4],

$$\Gamma_{X\to YZ} = \frac{|p|}{32\pi^2 m_X^2} \int |\mathcal{M}_{X\to YZ}|^2 d\Omega, \tag{1.7}$$

where $|\mathcal{M}_{X\to YZ}|$ is the scattering probability amplitude squared, $|p|$ is the modulus of the four-momentum of one[†] of the daughter particles, and $d\Omega$ is the solid angle. In case the mass of the daughter particles are negligible compared to the mass of the parent particle, the $|p|$ then simplifies to to $|p| \approx m_X/2$. With the definition of the decay rates, the branching ratio (or fraction) is defined to be,

$$\mathcal{B}(X \to YZ) = \Gamma_{X\to YZ}/\Gamma_{\text{total}}, \tag{1.8}$$

where Γ_{total} is the total decay rate of a given particle. It is simply the proportional fraction of the total decay rate.

It is interesting to note that the scattering probability amplitude for most SM interactions is $|\mathcal{M}_{X\to YZ}|^2 \propto g_c^2$ to leading order and this implies $\Gamma_{X\to YZ} \propto g_c^2$. Assuming the ignored portion of the $\Gamma_{X\to YZ}$ are similar for different decays, the sizes of branching ratios can be directly related to the coupling constant g_c of each decay. For the W or Z bosons, the coupling constants for each decay are the same size and hence the branching ratios are more or less democratic. For the Higgs boson, the coupling constants for each decays are different, where the larger the mass of the decay product, the larger the coupling constants. This leads to a hierarchical structure in the branching ratios with heavier particles taking up larger fractions of the total decay rate.

[†]It does not matter which daughter particle as the four-momentum are the same in rest frame of the parent.

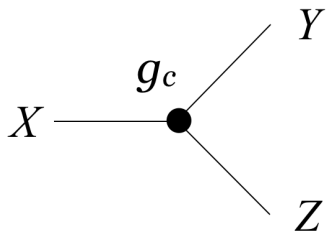


Figure 1.8: Illustration of a two-body decay of a parent particle X to daughter particles Y and Z with coupling constant of g_c.

1.8 Finding the Needle in a Haystack

For both the production cross section and the decay rate, depending on the kinematic phase-space choice, the sizes differ. For instance, consider a $Z \to \mu^+\mu^-$ process. The distribution of the invariant mass of the $\mu^+\mu^-$ system $m_{\mu\mu}$ has a resonance peak around the mass of the Z boson m_Z. If a phase-space requirement of $|m_{\mu\mu} - m_Z| > 15$ GeV is required the rate of the Z boson events will be drastically reduced. The basic strategy for particle searches is then to select a phase-space where the signal rate is maximized while the background rate is relatively low.

For each analysis, one often focuses on a set of final states. Once the final states are selected, there are two types of backgrounds: reducible and irreducible backgrounds. The reducible backgrounds are the type of background process where the final states are not identical to the focused final states; the irreducible backgrounds are the opposite. The reducible backgrounds often are heavily suppressed as soon as the data events are required to have the set of reconstructed final state particles. An example would be the reduction of W boson background when two the events are required to have two leptons (i.e. $\ell = e$ or μ) as W boson background only produces one ℓ. The irreducible backgrounds cannot be suppressed in the same way as the reducible backgrounds but only through kinematic selections. An example would be the reduction of $t\bar{t}$ background process by requiring the invariant mass of the two leptons $m_{\ell\ell}$ to be close to the Z boson mass m_Z. As the charged leptons from the $t\bar{t}$ decays do not need to form $m_{\ell\ell}$ close to the m_Z, the requirement of $|m_{\ell\ell} - m_Z| < \epsilon$ with some small ϵ will suppress $t\bar{t}$ backgrounds.

Typical data analyses go through several optimization process to select the optimal final states and kinematic selections to maximize signal while keeping the background rate low. One or more key variable distributions are fitted to make measurements, such as cross section, coupling strength, exclusion limit, or discovery significance.

Chapter 2

Physics of the Higgs Boson at the Large Hadron Collider

The Higgs boson in the SM explains how fundamental particles acquire mass through the Higgs mechanism [6–11]. The Higgs mechanism was proposed in the 1960s to explain how fundamental particles gained mass. Without the Higgs mechanism, mass terms of the gauge bosons and fermions are not allowed by the gauge symmetry $G_{EW} = SU(2)_L \otimes U(1)_Y$ of the SM, where G_{EW} is the electroweak symmetry governing the electroweak interaction. The Higgs mechanism spontaneously breaks the electroweak symmetry and allows mass terms of gauge bosons and fermions. This mechanism is called the electroweak symmetry breaking (EWSB). The consequence of the Higgs mechanism is an additional scalar particle the Higgs boson. The mass of the Higgs boson, then only a hypothetical particle, was not known and even some theorists even discouraged big experiments to search for the elusive Higgs boson [12].

In 2012, the discovery of a new boson at a mass of 125 GeV shook the field of particle physics. The next year physicists further continued to study the properties of the newly discovered boson. To date, all the measurement supports that the particle is consistent with the elusive Higgs boson predicted by the SM. As a result, the Nobel Prize in Physics in 2013 was awarded to Peter Higgs and Francois Englert for their early work on the theory of the Higgs boson.

2.1 Higgs Mechanism

The SM uses the language of quantum field theory to describe the fundamental particles and their interactions. The particles are the excitations of fields and their interactions are governed by the Lagrangian terms of the SM. The SM Lagrangian contains all possible terms allowed by the symmetries of nature with mass dimensions equal to four allowed by the symmetries of nature. The simplicity that a handful of symmetries of nature govern the dynamics of the particles makes the SM an *elegant* theory. Despite its elegance, the symmetries of nature does not allow mass terms for the bosons and fermions such as,

$$\frac{1}{2}m^2 A_\mu A^\mu, \quad m\bar{\psi}\psi, \tag{2.1}$$

where the A_μ is a boson field, and ψ is a fermion field. This can be seen by considering the gauge transformations of each field,

$$A_\mu \to A_\mu + \frac{1}{g}\partial_\mu \alpha, \tag{2.2}$$

$$\psi \to (1 + i\alpha)\psi, \tag{2.3}$$

where under such transformations the mass terms would not be invariant. Without breaking the symmetry explicitly one can *spontaneously* break the symmetry and generate mass terms if the ground state of the system does not obey the same symmetry. To see how this works, consider a complex scalar field ϕ with a potential,

$$V(\phi) = -\mu^2|\phi|^2 + \frac{\lambda}{2}|\phi|^4, \tag{2.4}$$

with specific requirement that $\mu^2 > 0$, which is the key for spontaneous breaking of the symmetry. Solving $\partial_\phi V(\phi) = 0$, one obtains that the expectation value of the lowest energy to be,

$$\phi_0 = \frac{\mu}{\sqrt{\lambda}}. \tag{2.5}$$

If $\phi_0 \neq 0$, the ground state energy is no longer invariant under the $U(1)$ transformation, while both the Lagrangian and the $V(\phi)$ is still invariant under the same $U(1)$ transformation. Such breaking of the symmetry is called the spontaneous symmetry breaking. If one expands ϕ around the ground state,

$$\phi = \phi_0 + \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2), \tag{2.6}$$

the kinetic term of the scalar field becomes,

$$|D_\mu\phi|^2 = \frac{1}{2}(\partial_\mu\phi_1)^2 + \frac{1}{2}(\partial_\mu\phi_2)^2 + \underbrace{g^2\phi_0^2 A_\mu A^\mu}_{\text{mass term for } A_\mu} + \underbrace{\sqrt{2}g\phi_0 A_\mu \partial^\mu \phi_2}_{\text{extra d.o.f for } A_\mu} + \cdots, \tag{2.7}$$

where several higher order terms are omitted and the mass term $g^2\phi_0^2 A_\mu A^\mu$ appears, which was not present before the spontaneous symmetry breaking. The additional direct coupling of gauge boson A_μ to ϕ_2 term allows ϕ_2 to act as an extra degree of freedom for the A_μ, allowing the gauge boson to acquire the necessary longitudinal polarization to be a massive particle.

The fermions gain masses after the spontaneous symmetry breaking in a different way. Mass terms for

the fermions are provided by the Yukawa interactions,

$$-\xi\phi\bar{\psi}\psi, \tag{2.8}$$

which becomes the fermion mass term after spontaneous symmetry breaking,

$$\underbrace{-\xi\phi_0\bar{\psi}\psi}_{\text{fermion mass term}} - \underbrace{\frac{1}{\sqrt{2}}\xi\phi_1\bar{\psi}\psi}_{\phi-\psi\text{ interaction}} + \cdots, \tag{2.9}$$

where the second term provides the interaction between the scalar and the fermion field.

If true, the Higgs mechanism can explain the massiveness of the SM particles. The consequence of such mechanism is the existence of a new massive scalar boson with the mass $m_{\phi_1} \equiv \sqrt{2}\mu = \sqrt{2\lambda}\phi_0$, which becomes apparent when the $V(\phi)$ is expanded to,

$$V(\phi) = -\frac{1}{2\lambda}\mu^4 + \underbrace{\frac{1}{2}m_{\phi_1}^2\phi_1^2}_{\phi_1\text{ mass term}} + \mathcal{O}(\phi_i^3). \tag{2.10}$$

2.2 Electroweak Symmetry Breaking and the Higgs Boson

So far the example focused on the abelian gauge field. Applying a similar mechanism to the electroweak symmetry, which is non-abelian, is known as the Glashow-Weinberg-Salam theory [13–15]. The associated scalar field the Higgs boson of the SM was searched for decades and was finally discovered in 2012. The discovery of the Higgs boson marks a triumph of the SM.

Before the discovery of the Higgs boson, it was not well understood how the vector boson scattering (VBS) process cross section was regulated. Fig. 2.1 shows the diagrams relevant for the VBS. The amplitude of the processes shown in the top row of Fig. 2.1 results in cross section rising as a function of center of mass energy $\hat{s}$ eventually breaking unitarity at $\hat{s} \geq 1 - 2$ TeV [17]. In SM, the diagrams in the bottom row of Fig. 2.1 exactly cancel the $\hat{s}$ dependence from the top row guaranteeing unitarity. If the Higgs boson were not discovered, the VBS process would have played a central role in understanding the EWSB. It might have led to discovering new physics that may explain some of the shortcomings of the SM with just one Higgs boson, such as the naturalness problem or the accomodation of fermion masses and mixing patterns [16].

Nature has chosen differently and the experiments are starting to see a more clear picture that the discovered Higgs boson does seem to play a role in unitarizing the VBS [18, 19]. However, an alternative mechanism may be at work in regularizing the VBS cross section. The couplings of the Higgs boson to the

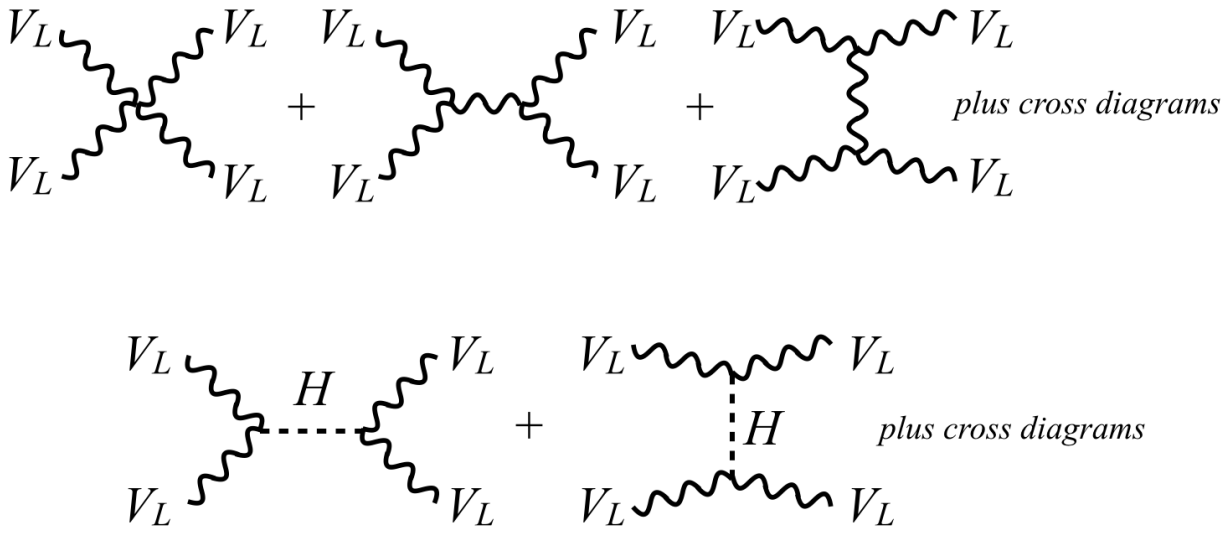


Figure 2.1: Feynman diagrams contributing to vector boson scaterring $V_LV_L \to V_LV_L$ process. The top row shows diagrams of vector boson scattering process with only vector bosons. The bottom row shows diagrams of vector boson scattering mediated by the Higgs boson.

gauge bosons may be altered compared to the SM predictions [20] or partial strong scattering [21] is at work. These possibilities would lead to deviations in VBS cross sections. The analysis presented in Chapter 5 serves as a step towards answering the puzzle.

2.3 Higgs Boson Couplings

The predicted SM Higgs boson couples to all other SM particles that have mass. The nature of the coupling between the Higgs boson to massive gauge boson is different from that of the fermions; as outlined in the previous section, the bosonic couplings arise from the EWSB process, while the fermionic couplings are from the Yukawa interactions. Although coupling strengths between the Higgs boson to different particles are different, they can be organized in terms of the mass of the particle the Higgs boson couples to. The couplings of the Higgs boson to fermions and vector bosons are [4],

$$g_f = \sqrt{2}\frac{m_f}{v}, \quad g_V = 2\frac{m_V^2}{v}, \tag{2.11}$$

and this implies that the g_f and $\sqrt{g_V/v}$ have a linear relation to m_f or m_V, where v is the vacuum expectation value of the scalar field responsible for the EWSB, and m_f and m_V are the masses of the SM fermions and vector bosons, respectively. This is illustrated in Fig. 2.2, where the coupling strengths show a linear relation to the mass.

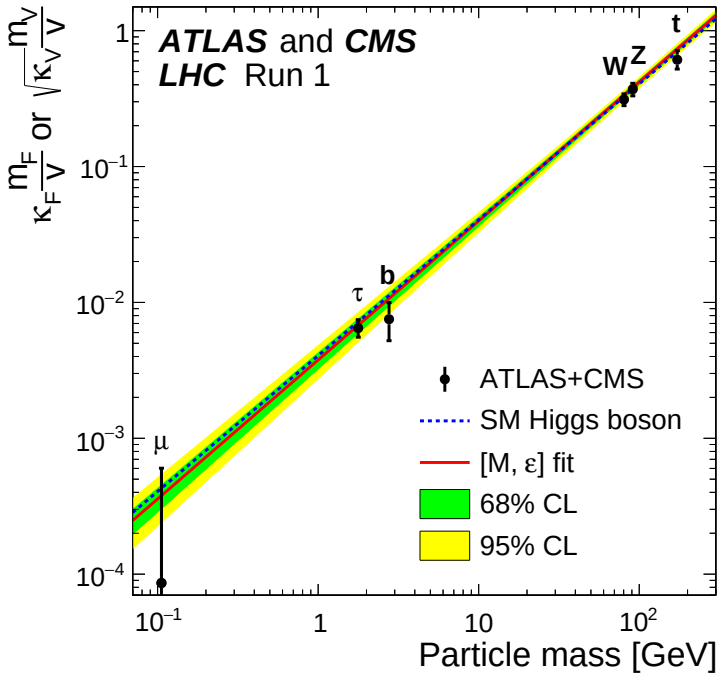


Figure 2.2: Linearity of Higgs boson couplings to other particles [22]. The Higgs boson coupling strengths are plotted against the mass of the particle the Higgs boson couples to. See texts for the discussion.

Precise measurements of the Higgs couplings are essential in confirming the nature of the newly discovered Higgs boson. If any of the couplings deviate from the expected linear relation, it would indicate a hint of new physics beyond the SM [23, 24]. The measurements presented in this thesis probe the Higgs couplings and in Chapter 7, constraints on the two different types of couplings will be presented.

2.4 Cross section and Branching Ratio of the Higgs Boson

The production of the Higgs boson through the pp collisions can occur in a number of ways, but only a few Higgs boson production process has a large enough cross section to be relevant for the pp collisions at the LHC. Feynman diagrams of the three leading production process are shown in Fig. 2.3. In these diagrams, the produced Higgs bosons decay to WW^*, which is the final state the thesis focuses on in Chapter 5. The cross section for these processes as a function of the Higgs boson mass is shown in Fig. 2.4. The gluon-gluon Fusion (ggF) process has the largest cross section by about an order of magnitude to the subleading vector boson fusion (VBF) process for most of the Higgs boson mass range. The vector boson associated (VH) process can be broken up into two separate processes where the vector boson can be either a W or a Z boson. The VH process cross section is few factors below the VBF process cross section.

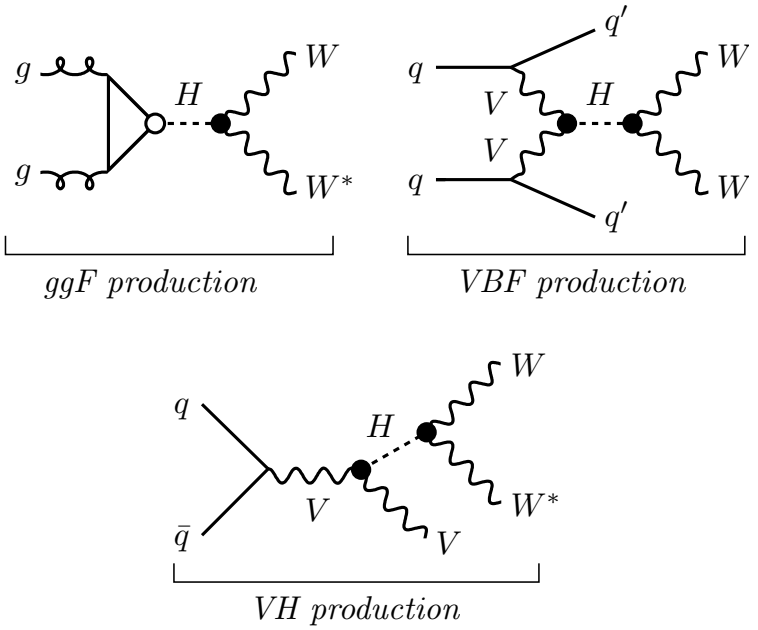


Figure 2.3: The illustration and the caption is from Ref. [18]. Feynman diagrams for the leading production modes (ggF, VBF, and VH), where the VVH and qqH coupling vertices are marked by • and ∘, respectively. The V represents a W or Z vector boson.

In Fig. 2.3, the Higgs boson production vertex and decay vertex are indicated by a solid or a hollow circle to indicate its coupling type. A • indicates a bosonic coupling, and a ∘ indicates a fermionic coupling as explained in the Sec. 2.3. In order to probe different types of couplings comprehensively, the Higgs boson search program at the LHC experiments looks for each process individually.

Before the discovery of the Higgs boson, the focus of the Higgs boson search was heavily on ggF process, which has the highest production cross section. After the discovery, the focus has been expanded to include other production mechanisms to characterize the newly discovered particle. This thesis documents the search for the Higgs boson with the subdominant VBF production process. It provides an independent information of the properties of the Higgs boson compared to the ggF process as the couplings involved in the production is different from that of the leading ggF process.

The Higgs boson is an unstable particle and thus decays to other SM particles. Fig. 2.5 shows the branching ratios of the Higgs boson to various other final states as a function of the Higgs boson mass. The discovered Higgs boson has a mass around $m_H = 125$ GeV. At this m_H value, Higgs bosons decay predominantly to b-quark pairs ($b\bar{b}$). The decay channels such as WW^*, ZZ^*, and $t\bar{t}$ which are naively expected to have high branching ratios are suppressed due to the fact that the $m_H < 2m_i$, where m_i is the mass of the final state particle. The particle with the highest mass that satisfies $2m_i < m_H$ is the bottom-quark; hence the branching ratio to $b\bar{b}$ final state is the largest.

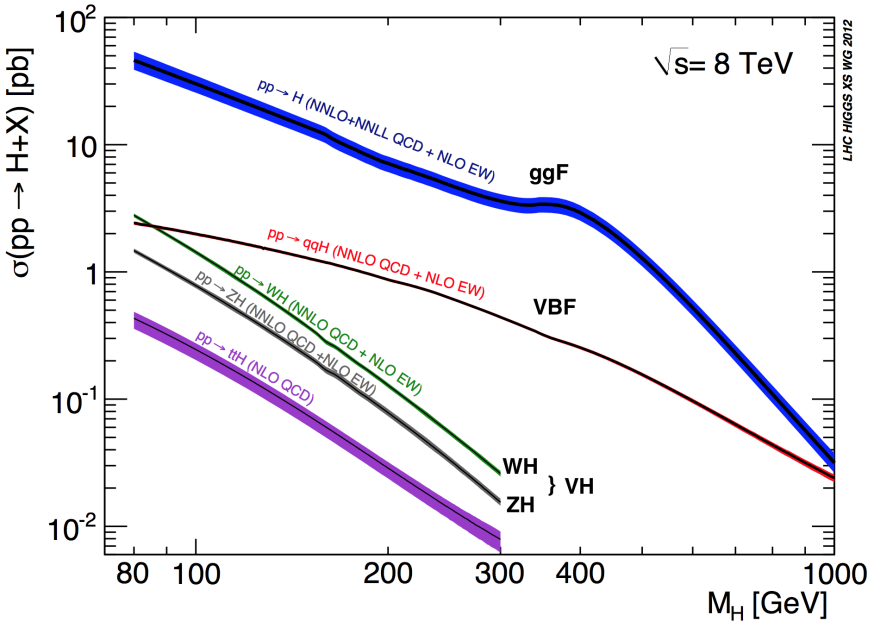


Figure 2.4: Standard Model Higgs Boson Production cross section at $\sqrt{s} = 8$ TeV. The cross section for the gluon-gluon Fusion process is the largest. The vector boson fusion (VBF) process has the second largest cross section [25].

The next largest branching ratio after $b\bar{b}$ is to the WW^* final state, where one of the W are expected to be off-shell and this thesis focuses on the WW^* decay channel. More specifically, the search is over the fully leptonic decay channel of the $WW^* \to \ell\nu\ell\nu$ final state, where $\ell \equiv e$ or μ.

It is important to note that a significant portion of the Higgs boson decays to final states with massless particles (i.e. gg, $\gamma\gamma$, or $Z\gamma$). As the Higgs boson does not couple directly to massless particles, these processes only occur through a loop diagram. Whenever loops are included in a SM process, it opens up to a possibility of non-SM particles in the loop altering the behavior of the process. The Higgs boson decay channel $H \to \gamma\gamma$ and the ggF process thus are good probes for non-SM effects in the Higgs sector.

2.5 Extended Higgs Sector Beyond the Standard Model

Implications of the Higgs boson discovery includes that nature contains a fundamental scalar particle. Fundamental scalar fields show up frequently in various models of physics beyond the SM (BSM) [26]. Although there can be no logical correlation between the prevalence of scalar fields in BSM theory literatures to the likeliness of nature containing extra scalar fields, the discovery of a fundamental scalar field naturally raises the question whether there can be more scalar fields in nature. The possibility of an *extended* Higgs sector is also realized in one of the popular BSM model the Minimal Supersymmetric Standard Model [27], where

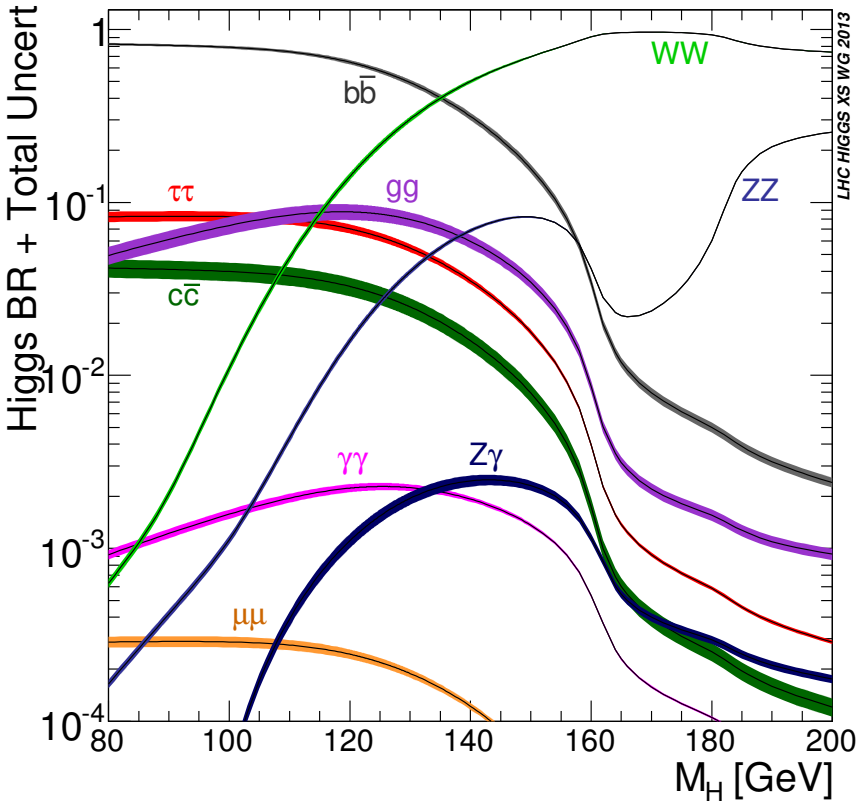


Figure 2.5: Standard Model Higgs Boson decay branching ratios as a function of the Higgs boson mass M_H. For $M_H = 125$ GeV, the largest branching ratio is to the $b\bar{b}$. The second largest is to the WW^*, which this thesis focuses on [25].

it contains more than one scalar fields.

The scalar fields are quite special. Consider the SM Higgs boson mass term in the SM,

$$-\mu^2 H^\dagger H, \tag{2.12}$$

which is the only term in the SM that does not have a mass dimension of four. The Higgs mass term has the coupling term μ^2 with a mass dimension of two and the operator $H^\dagger H$ with a mass dimension of two. The mass dimension being smaller than two means that the Higgs mass term is open to renormalizable couplings to other scalars including SM singlet fields [28]. If there were an additional scalar s in nature, an interaction such as,

$$\frac{\xi}{2} s^2 |H|^2, \tag{2.13}$$

can provide a sizable branching ratio of $\mathcal{B}(H \to ss) = 10\%$ even for a coupling strength as small as $\xi = 10^{-2}$ [30]. The Higgs boson is thus a prime candidate to search for a non-SM decay to extra scalars. In light of this, searches for $H \to ss$ are being conducted at the LHC. One of such search is documented in Chapter 6.

Chapter 3

ATLAS Experiment

3.1 Detector

The general-purpose detector ATLAS is one of the four major detectors at the LHC [31]. The ATLAS detector is forward-backward symmetric and is in a cylindrical shape that has near 4π coverage in solid angles. A schematic picture of ATLAS detector is shown in Fig 3.1.

The inner detector (ID) consists of three subdetectors that detect as the charged particles traverses through the detector. The ID has a coverage of $|\eta| < 2.5$. The innermost subdetector is the pixel detector. The pixel detector consists of four layers of pixel detector and tracks the discrete point that the charged particle traversed. The innermost layer is the Insertable B-Layer (IBL) pixel detector which has been added during the long shutdown 1 (LS1). The pixel detector closest to the interaction point is most useful in finding relatively long-lived particles such as hadrons with b-quarks; for this reason, it is referred to as the "b-layer". During the 2011 and 2012 data taking, the IBL was not present. The next subdetector part is the SemiConductor Tracker (SCT). The SCT consists of four double layers of strip detectors, which behaves similarly to the pixel detector except the measurement is one dimensional. The last tracker layer is the transition radiation tracker (TRT), which consists of xenon gas filled tubes and wires. The superconducting solenoid magnet applies magnetic field with a strength of 2 T, bending the charged particles' path depending on their momentum.

The calorimeter system surrounds the inner detector and the solenoid magnet. The total coverage of the calorimeter system is up to $|\eta| < 4.9$, while each subdetector has different $|\eta|$ coverage. The calorimeter system consists of three parts: electromagnetic (EM) calorimeter, hadronic calorimeter, and forward calorimeter (FCal). The calorimeter subdetectors force all particles to release their energy into the calorimeter through particle-material interaction referred to as "showers". The energy deposits "activate" either the scintillator or the ionization medium and subsequently the "activity" is read out through electronics. The granularity of the electronic readout in the calorimeters provide spatial information about the energy deposits, which allows for particle spatial reconstruction.

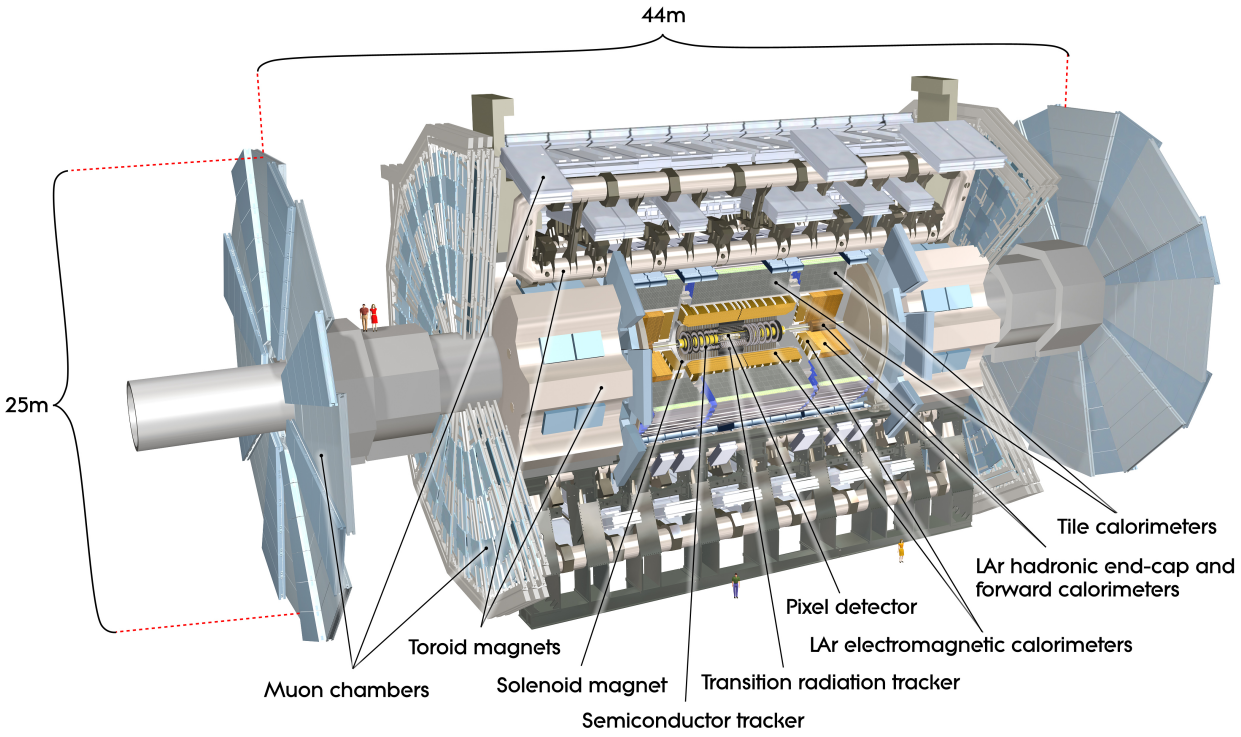


Figure 3.1: ATLAS detector is shown with each subdetector parts labeled (ATLAS Experiment © 2017 CERN).

Two types of calorimeters are used in the ATLAS detector. The liquid-argon (LAr) calorimeter uses lead as stopping material and LAr as sampling ionization medium to read out the amount of energy deposited. The tile calorimeter (TileCal) uses steel as stopping material and scintillators to read out the amount of energy deposited. The LAr calorimeter technology is used for the EM calorimeter and provides a coverage of $|\eta| < 3.2$ and measures electromagnetic showers. The same LAr calorimeter technology is used for hadronic calorimeter in the region of $1.5 < |\eta| < 3.2$. The TileCal technology is used in the central region $|\eta| < 1.7$. The FCal covering the region of $3.1 < |\eta| < 4.9$ consists of only the LAr calorimeters for the measurements of both electromagnetic and hadronic showers.

The muon spectrometer (MS) is designed to detect muons in the region of $|\eta| < 2.7$. The MS surrounds the calorimeters and sits at the outermost layer of the entire detector. The barrel region of the MS is defined to be in the region of $|\eta| < 1.05$, while the endcap regions cover the rest. The MS is supplemented by the three large toroid magnets, each with eight coils, providing a magnetic field with a bending power of about 2.5 $\mathrm{T} \cdot \mathrm{m}$ in the barrel region and up to 6 $\mathrm{T} \cdot \mathrm{m}$ in the endcaps. Monitored drift tube chambers covering most of the η range in both the barrel and endcap regions along with cathode strip chambers covering $2.0 < |\eta| < 2.7$ are used as precision-measurement chambers, whereas resistive plate chambers (RPC) in the barrel and thin gap chambers (TGC) in the endcaps are used as trigger chambers, covering $|\eta| < 2.4$. Three layers are arranged in the chambers, so a coincidence requirement can be made.

3.2 Trigger and Data Acquisition

The *pp* collisions occur at a rate of 40 MHz. The data size of each event is around 1.6 MB. If every events are recorded, this would amount to a data flow of 52 TBps. To put this number in perspective, consider the average amount of data uploaded to the internet, which is estimated to be 28 TBps [32]; the ATLAS detector is producing twice the amount of data the entire internet is using over the globe. It is therefore unrealistic to record every events being produced by the ATLAS.

The "trigger" system solves this problem by only recording interesting events. A multi-staged trigger system quickly sifts through the event in real time and makes the decision whether to save the data or not. The first stage is called the level-1 (L1) trigger and it uses dedicated hardwares. The second stage is called level-2 (L2) trigger and the last stage is called the event filter (EF). Both L2 trigger and EF performs decision making process on a dedicated CPU farm. The first stage (i.e. level-1 or L1) uses reduced granularity information to quickly decide what objects are present in the event and whether the event is worth passing to the next stage of the trigger system. At this stage, the RPC and TGC in the MS performs

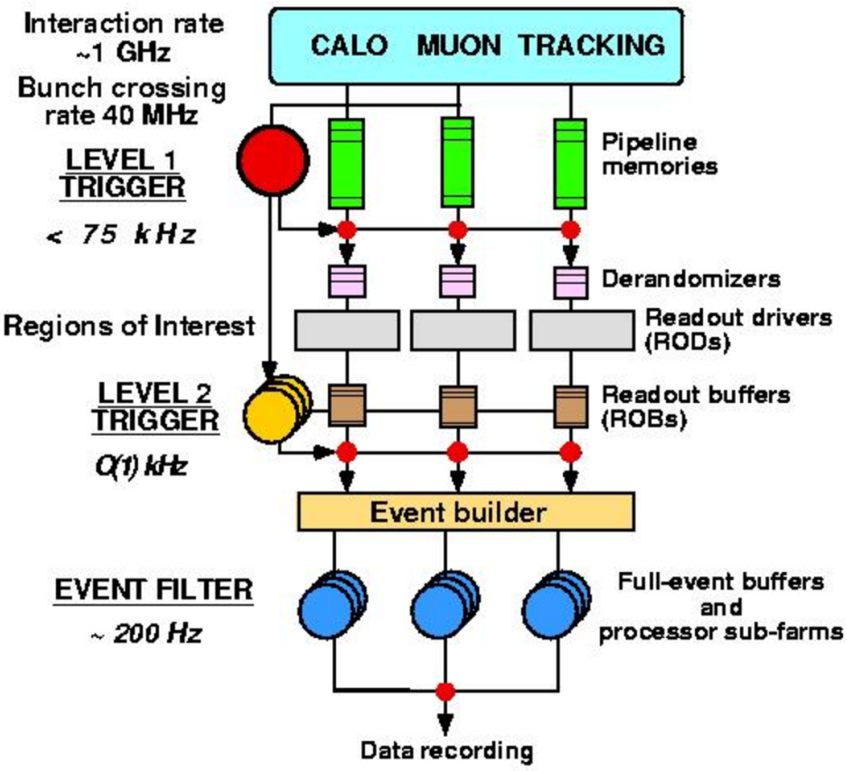


Figure 3.2: Overview of the trigger system used in Run 1. Three stage trigger is used. The Level 1 trigger reduces the rate from 40 MHz of bunch crossing to around 75 kHz. The Level 2 trigger further reduces the rate to $\mathcal{O}(1)$ kHz. The Event Filter (EF) reduces the rate down to $\mathcal{O}(100)$ Hz [33].

a quick scan to look for coincidences to find high-p_{T} muons. The calorimeter granularities are reduced to bigger cells and a quick 3×3 scan is performed to look for high-p_{T} objects. At the output of the L1 trigger, the rate is reduced from 40 MHz to 75 kHz, which corresponds to factor of 500 reduction in rate.

If the event has passed the L1 trigger, the read out electronics retrieve the information from the detector and save the event data into buffers. The L2 trigger focuses on the "Region of interests" (ROIs) provided by the L1 trigger and uses the full information in the subregion of the detector to refine the object selection with higher resolution than what was used in L1 trigger. At the output of the L2 trigger, the rate is reduced from 75 kHz to 3.5 kHz, which corresponds to factor of 20 reduction in rate.

The last stage of the trigger is the EF, which utilizes the full ATLAS reconstruction of objects and perform event selection based on high resolution objects. The final output event rate is of $\mathcal{O}(100)$ Hz. The overview of the trigger system is shown in Fig. 3.2 for the 2011 and 2012 data taking.

For the 2015 and 2016 data taking, the trigger system underwent upgrades during the LS1 [34]. Several aspects of the hardware L1 trigger were upgraded to have better resolution. The L2 and EF were combined

into a single stage called high-level trigger (HLT). As a result the data acceptance rate increased to 1 kHz from $\mathcal{O}(100)$ Hz.

3.2.1 Fast TracKer (FTK) Upgrade

Tracking information in hadron colliders are essential for identifying particles accurately. Track is reconstructed from inner detector traces. As the inner detector only provides points in space where a charged particle has gone through, the reconstruction step involves a large number of combinatorics to consider. This in turns leads to intensive usage of computing resources.

Once a data event is accepted by the L1 trigger, the HLT will begin to process the data. At the beginning of this stage, no tracking information is available. The HLT system will have to use its resources to reconstruct tracks by considering a large number of combinatorics. To help HLT system, the FTK system will use dedicated hardware to reconstruct tracks and provide tracking information to the HLT with low latency [35]. With the help of tracking information, the data reduction rate will increase at the HLT, allowing for a larger phase-space acceptance by loosening event selection criteria.

The expected improved trigger performance with the usage of FTK for b-jet trigger is shown in Fig. 3.3. The x-axis shows the efficiency for signal events and y-axis shows the output event rate. The red dots, which are the triggers using FTK tracks, are shown to be on the right handside of the black dots, which are the triggers not using FTK tracks. The plot shows that for a same amount of output event rate more signal events are being accepted.

The FTK system operates through several stages. The first stage is the read out stage. Once the L1 trigger accepts an event, the FTK system starts to read out both the pixel and SCT detectors. It then undergoes a clustering process where the energy traces of charged particles are clustered to form *hits*. With the usage of Associative Memory chips, which performs fast pattern matching, track candidates are formed using only the 8 out of 12 layers of trackers. Track candidates are sent to the *second stage fit* (SSF), in which the rest of the four layers of informations are combined to form a full 12-layer tracks. The Illinois group led by my advisor took the responsibility of building the *second stage board* (SSB), which performs the SSF. The Fig. 3.4 shows a picture of SSB.

My role in the development of the SSB included development of the firmware responsible for extrapolating tracks from the first stage fit to the additional four layers and aggregating list of hits to be associated with a track. I have also taken the role of writing the software simulating the SSF for the official ATLAS simulation. In the process, studies were done to optimize the SSF algorithm. Lastly, I have been involved with commissioning of SSBs and integrating the FTK into the ATLAS trigger system.

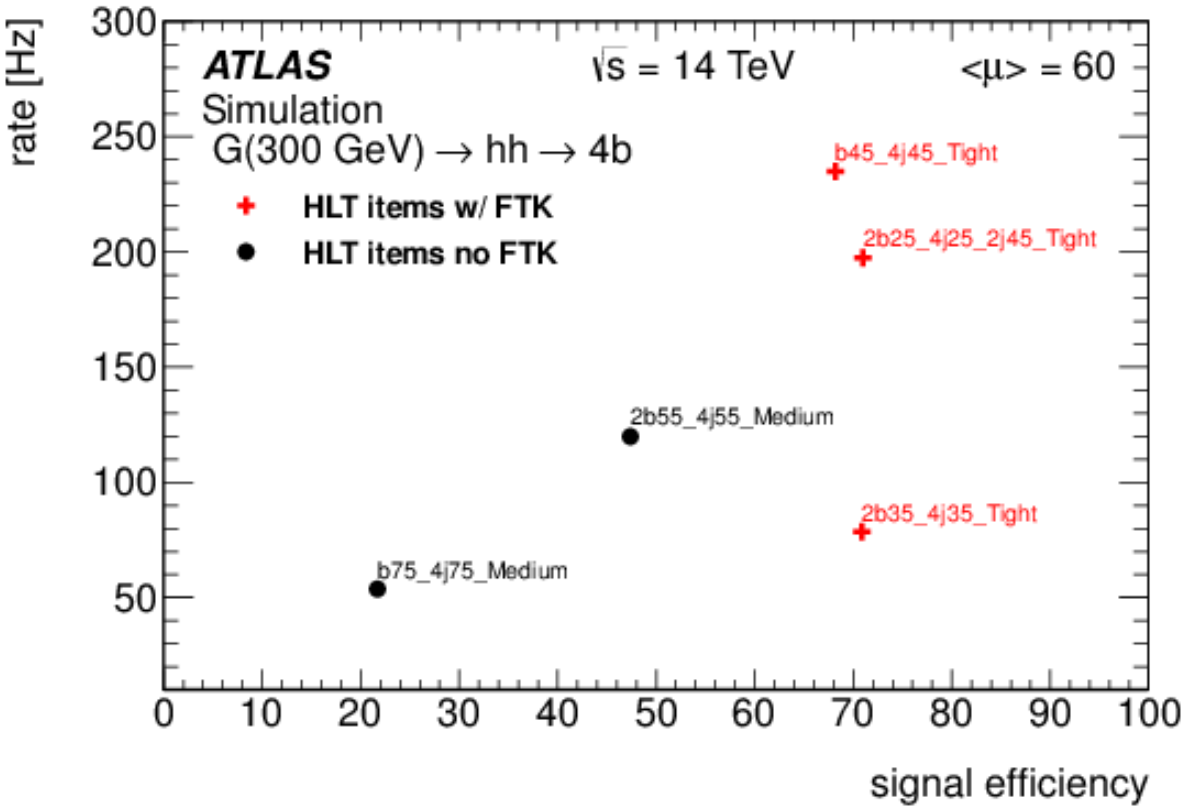


Figure 3.3: Comparison of trigger configurations with and without FTK is shown. The x-axis shows the efficiency of accepting signal events for a trigger. The y-axis shows the trigger rate for a given trigger configuration. An ideal trigger configuration will have a high signal efficiency with a low trigger rate. The black (red) dots show the trigger configuration without (with) FTK. The plot shows that with a similar trigger rate, signal efficiency can be higher for trigger configurations using FTK [36].

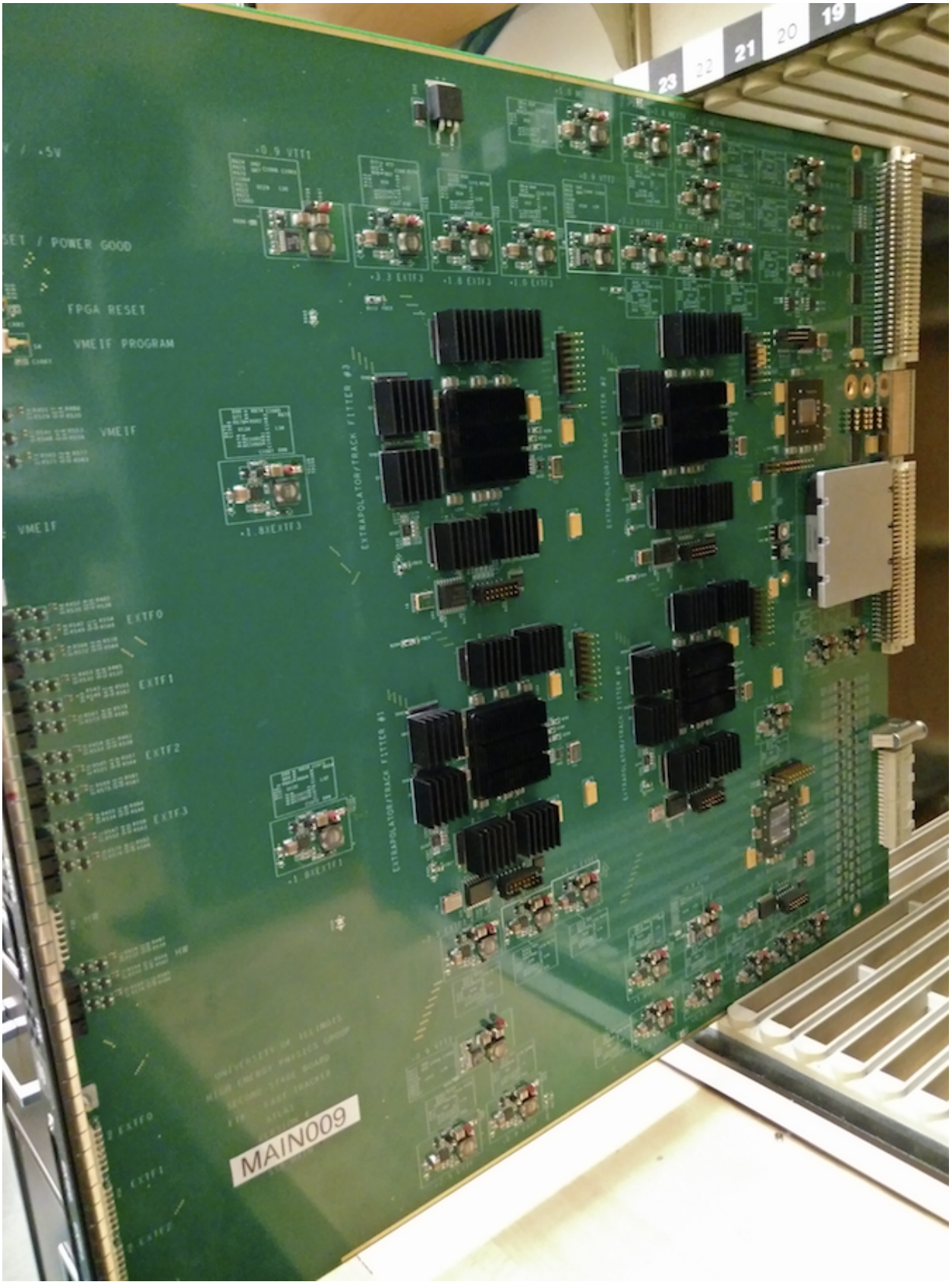


Figure 3.4: One Second Stage Board of the Fast TracKer system.

3.3 Computing

The vast amount of data ATLAS detector produces are analyzed by physicists around the globe. The Worldwide LHC Computing Grid (WLCG) is used to store, distribute, and analyze data for physics analysis. The WLCG consists of four levels, Tier 0 to Tier 3. The Tier 0 is the CERN's local data center. The event data recorded from the ATLAS detector is in the format named RAW. Every event recorded by the ATLAS detector first gets processed at Tier 0. The Tier 0 processes the RAW data and performs event reconstruction on all data. Then the processed data and the original RAW data are distributed to various Tier 1 sites.

The Tier 1's consists of 13 sites, which usually are national labs or computing center. Each Tier 1 sites have large enough storage to copy event data from Tier 0 to their respective sites. Each Tier 1 sites are linked to Tier 0 site over LHC Optical Private Network, which provides the connection between Tier 0 to Tier 1s with 10 Gbps.

The Tier 2 facilities are typically institutes that can provide sufficient storage and computing resources. There are about 150 Tier 2 facilities around the world. Most analysis tasks are performed in Tier 2 facilities.

Lastly the Tier 3 consist of local clusters at each university department and can be used to access the grid. Individual computers can be used to directly access the grid as well.

Chapter 4

Experimental Methods

Up to this chapter, the aspects of how the pp collisions occur are discussed and how the data is taken. The data taken from the detector are in RAW format (i.e. ones and zeroes). To make sense of the data, one needs to process them into sensible objects. The process of inferring a physics object (e.g. electron, muon, τ, jets and etc.) from electrical signals of the detector is called the reconstruction. Once candidate objects are identified various kinematic quantities are measured.

As for most experiments, the measured quantities need to be compared to the expectations. Often in hadron collider physics, obtaining the expected distributions through exact calculations is difficult for few reasons viz. QCD is non-perturbative, calculations for $\mathcal{O}(100)$ particles computationally infeasible. Expected distributions of the pp collisions are thus modeled through Monte Carlo (MC) methods, which are more tractable.

4.1 Event Reconstruction

The detector measures low-level primitive quantities such as clusters of hits, energy deposits, segments of muon tracks and etc. From the primitive quantities one defines the objects. The reconstructed objects *s.s.* are not the real physics objects, so there are possibilities of other physics objects being misidentified, or worse, random noises in the detector faking an object. The reconstruction thus focuses on two quantities: the particle identification (PID) efficiency, and the fake rate. For a tight identification of an object, the fake rate is kept at a low value in exchange for a loss in PID efficiency and vice versa for loose identification. A medium identification makes a compromise and keeps the efficiency and the fake rate somewhere in between. In practice, several working points are provided for a given object definition, varying in PID efficiency ϵ and the corresponding fake rate f_r. The Fig. 4.1 shows an illustration of various working points on the receiver operating characteristics (ROC) curve. The ROC curve shows ϵ on x-axis and f_r^{-1} on y-axis. As the identification criteria are tightened, the PID efficiency goes down while the inverse fake rate increases. The ROC curve of a hypothetical alternative PID is shown in red curve. The alternative PID is better for

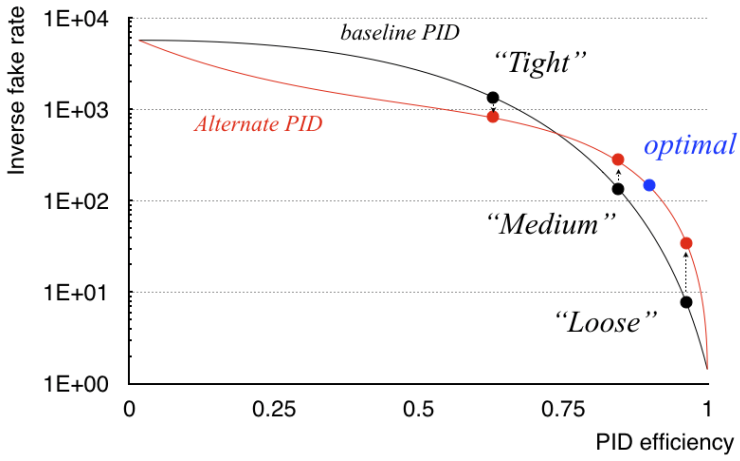


Figure 4.1: Cartoon of ROC curve for object identification. The x-axis shows the efficiency of a given particle identification. The y-axis shows the rejection power of misidentified or faked objects. The cartoon illustrates several different working points for a given particle identification algorithm. An alternative particle identification algorithm is also shown to illustrate a scenario of comparing two different particle identification algorithm. Several working points for the two different particle identification algorithm are shown. The optimization process for an analysis would consider several multiple working points and choose the optimal point, as illustrated by the blue dot.

given signal efficiencies corresponding to *medium* and *loose* but not for *tight* PID. The optimization process for an analysis considers several multiple working points and chooses the optimal point, as illustrated by the blue dot.

4.1.1 Muons

Muons are reconstructed by matching an ID track segment with a segment in the MS detector. The MS tracks must traverse two out of the three layers. A similar quality requirement is applied to the ID tracks, where at least a minimum number of hits associated for each ID subdetectors are present with each ID track. For the VBF analysis documented in Chapter 5, the muon candidates referred to as combined muons in Ref. [37] are used, where the track parameters of the MS track and the ID track are combined statistically. Muon candidates are required to have $|\eta| < 2.5$. The detailed description of efficiencies and the fake rate is provided in Ref. [37].

4.1.2 Electrons

Electrons are reconstructed by matching the cluster of energy deposited in the EM calorimeter with tracks from the ID [38]. All electron tracks used in the analysis for the Chapter 5 uses a fitting method called Gaussian sum filter [39] (GSF) to account for bremsstrahlung energy losses. The difference between the

clustered energy deposit in the calorimeter and the track momentum measured by the ID are reduced by the GSF fit, which in turn improves the angular resolution and impact parameter* resolutions. As the calorimeters measure the energy deposits for the electrons, the four-momentum of electrons are reconstructed through its energy measurement.

The electron candidates are required to have $|\eta| < 2.47$, excluding the region between the barrel and endcap EM calorimeters, $1.37 < |\eta| < 1.52$. Several key signatures of electrons in each part of the detectors are used to identify electron candidates. The electromagnetic shower shapes in the EM calorimeters are required to be consistent with what is expected from an electron, the EM clusters and the ID tracks need to match in $\eta - \phi$ space, and signals of transition radiation are required in the TRT. The efficiency of electron identification using these characteristics is improved with the adoption of a likelihood-base method in addition to the default selection-based method. The selection-based method incurs a loss of efficiency when rectangular cuts are made in the high dimensional space. The likelihood-based method selection instead correctly carves out the phase-space for an improved efficiency while keeping the fake rate low. Detailed discussions of the electron identifications and the corresponding efficiency measurements can be found in Ref. [40]. Electrons with $10 < E_{\mathrm{T}} < 25$ GeV are required to satisfy the "very tight" likelihood-base identification. The fake contribution from jets and photons being misidentified as an electron is reduced by 35% when using "very tight" likelihood-base identification relative to the selection-based identification for the same signal efficiency. For $E_{\mathrm{T}} > 25$ GeV the electrons are required to satisfy the "medium" selection-based identification, where the fake rate is less concerning.

4.1.3 Jets

Particles carrying color charges does not exist in isolation due to what is known as color confinement. Once produced, quarks and gluons *hadronize* by pulling other quarks from the vacuum; the hadrons decay into lighter SM particles producing a *jet* of particles. The energy deposits from a jet is measured by the detector. Through clustering of the energy deposits a jet is reconstructed, and a presence of a quark or a gluon from the pp collision is inferred.

Jets are reconstructed using the anti-k_t clustering algorithm [41,42] with a radius parameter $R = 0.4$. The clustering algorithm takes three-dimensional clusters of energy deposited in the calorimeters as inputs [43,44]. The anti-k_t clustering algorithm is effective against noise by clustering significant energy clusters and their neighboring ones. Prior to the clustering, each cluster is calibrated based on whether the cluster is from an

*The impact parameter is the lepton track's distance of closest approach in the transverse plane to the reconstructed position of the primary vertex. The electron transverse energy is computed from the cluster energy and the track direction at the interaction point.

electromagnetic or a hadronic shower [45]. An additional correction is applied to each cluster to mitigate contributions from in-time and out-of-time pile-up [46]. The reconstructed four-momentum of jets are calibrated to the correct hadronic energy scale primarily determined by the simulation and further corrected by correction factors from data [44, 45]. The systematic uncertainties are assigned for the calibrations and the details can be found in Ref. [44, 45].

As jet reconstructions take clusters of energy from calorimeter as inputs, they are susceptible to the noise contributions from pile-up interactions. The larger the number of pile-up interactions, the more jets are produced in each bunch crossing. The presence of a large number of jet candidates from pile-up vertices potentially affect each analysis result if the pile-up jet is selected instead of the correct jet from the primary vertex of interest. To reduce jet candidates originating from pile-up vertices, the pinpointing capability of ID tracks is used to associate each jet with its originating primary vertex. A quantity called jet vertex fraction (JVF) is calculated, where the JVF is defined as the fraction of summed scalar p_{T} of tracks inside the jet radius from the associated primary vertex. A requirement on JVF is made for each jet with $p_{\mathrm{T}} < 50$ GeV [47], since pile-up jets on average has low p_{T}. No JVF selection is required for jets with no associated tracks.

4.1.4 b-tagging

The jets originating from b-quarks tend to form secondary vertices away from the primary vertices. The presence of a secondary vertex inferred from several tracks with high impact parameter is a distinct signature of b-quark orginated jets. Several quantities related to the characteristics of b-quark originated jets are fed into multivariate technique—the MV1 algorithm [48, 49]—to discriminate b-quark (or c-quark) originated jets from "light" jets, which originates from u, d, s-quarks or gluons. To tag a jet as "b-jet", a selection is required on the output of the MV1 algorithm, which is a continuous distribution ranging from -1 to 1 with higher values signifing that the jet is more likely to be originated from a b-quark. As the tracker plays a key role identifying secondary vertices, b-tagging is limited to jets with $|\eta| < 2.5$. Several working points are provided and calibrated for b-tagging and corresponding efficiencies are discussed in detail in the Ref. [50].

4.1.5 Overlap Removal

Any pair amongst lepton and jet candidates may be close together in $\eta - \phi$ space. If a pair has candidate objects geometrically overlapping significantly, following procedures are applied to remove one of the overlapping objects. If a muon candidate and an electron candidate are separated by $\Delta R < 0.1$, then the muon is retained, and the electron is removed. These cases usually indicate a muon that has undergone bremsstrahlung in the ID material or calorimeter. A high-p_{T} electron is almost always also reconstructed

as a jet, so if an electron and the nearest jet are separated by less than $\Delta R = 0.3$, the jet is removed. A muon candidate may arise from a heavy-flavor decay of a jet, so if a muon and a jet overlap with less than $\Delta R = 0.3$, the muon candidate is removed. Due to an early bremsstrahlung, a prompt electron may produce more than one electron candidate in its vicinity. In the case of two electrons separated by less than $\Delta R = 0.1$, the electron candidate with larger E_{T} is retained.

4.1.6 MET

If a high-p_{T} neutrino is present in the event, the signature shows up as a momentum imbalance in the transverse plane. The momentum imbalance calculation be summarized as

$$\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}} = -\left(\sum_{\mathrm{selected}} \boldsymbol{p}_{\mathrm{T}} + \sum_{\mathrm{soft}} \boldsymbol{p}_{\mathrm{T}} \right), \tag{4.1}$$

where the selected refers to the objects reconstructed (viz. leptons, photons, and jets), and the soft refers to remaining low-p_{T} objects not identified. Different methods of evaluating the soft objects result in different definitions of the missing transverse momentum.

The standard approach is to use the calorimeter-based reconstruction. The calorimeters provide a large coverage in η and are sensitive to neutral particles. In addition to the selected objects, jets and photons with $E_{\mathrm{T}} > 20$ GeV are vectorially summed over. The remaining soft objects are identified by the net transverse momentum of the calibrated calorimeter cluster energy measurements. The resulting missing transverse momentum is denoted $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$.

The resolution of the calorimeter-based measurement of missing transverse momentum suffers when significant pile-up is present. An alternative method of using a track-based measurement of the soft objects improves the resolution by $\mathcal{O}(20\%)$. The tracks are required to have $p_{\mathrm{T}} > 0.5$ GeV and originate from the primary vertex. To avoid double counting when estimating the soft objects, tracks associated with selected leptons or jets are not included. This reconstruction of missing transverse momentum, denoted $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{miss}}$, is used in the calculation of various variables using the missing transverse momentum.

Fig. 4.2 shows the expected resolution for the magnitude of $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ and $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{miss}}$ ($E_{\mathrm{T}}^{\mathrm{miss}}$ and $p_{\mathrm{T}}^{\mathrm{miss}}$, respectively). The r.m.s. of the missing transverse momentum difference decreases from 15.9 GeV to 12.4 GeV when using $p_{\mathrm{T}}^{\mathrm{miss}}$ instead of $E_{\mathrm{T}}^{\mathrm{miss}}$ in the reconstruction.

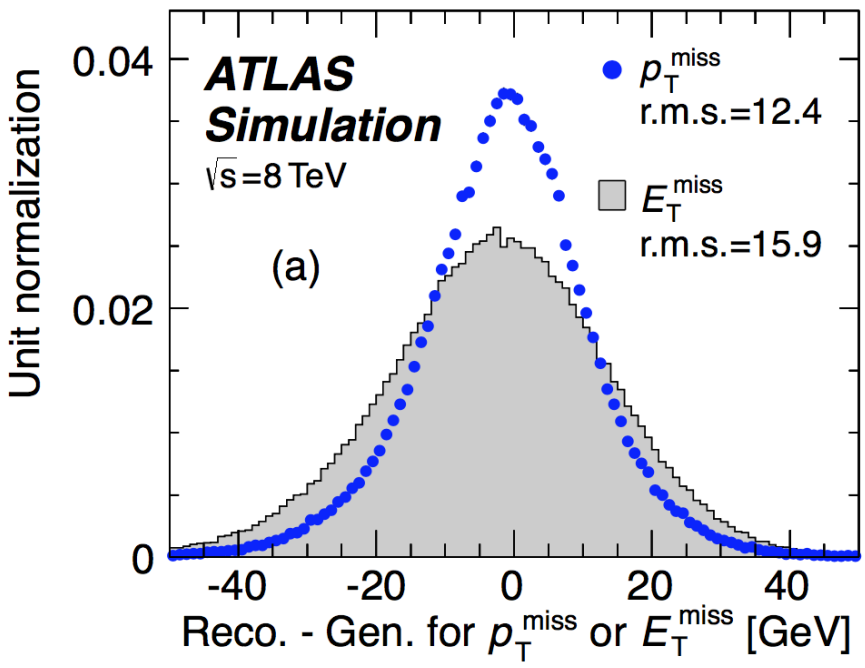


Figure 4.2: Comparison of the resolutions of missing transverse momentum for the calorimeter-based reconstruction ($E_{\mathrm{T}}^{\mathrm{miss}}$) and the track-based reconstruction ($p_{\mathrm{T}}^{\mathrm{miss}}$) of the soft objects [18].

4.2 Monte Carlo Simulation of pp Collisions

The pp collision data are compared to the simulated pp collisions. Each collision events involve several hundreds of particles, and an exact calculation of the dynamics is nearly impossible. Even if the large number of particles can be handled in simulation, the solution to QCD is not well understood. A MC method is adopted in simulating the events instead.

The MC generators will sample pp collision events from the approximate model of a given process and produce a large sample of simulated pp collision events. The produced events will contain detailed descriptions of all particles in the event; the detailed description is referred to as the *truth* information. The particle-detector interactions are simulated as well and the expected digital signals through the detector for the simulated pp collision events are computed. The digitized information is stored as the same RAW format the data events are in. Identical reconstruction methods used during the processing of collision data events are applied to the MC sample events.

Several MC programs are available. The schematic shown in Fig. 4.3 summarizes the simulation processes. The simulation of a pp collision starts with describing the partons colliding from each of the proton. Again, several choices of PDFs are available. The incoming partons to outgoing parton scattering processes are modeld by the scattering matrix $|\mathcal{M}|^2$ computed from Feynman rules of the SM by various generators. The

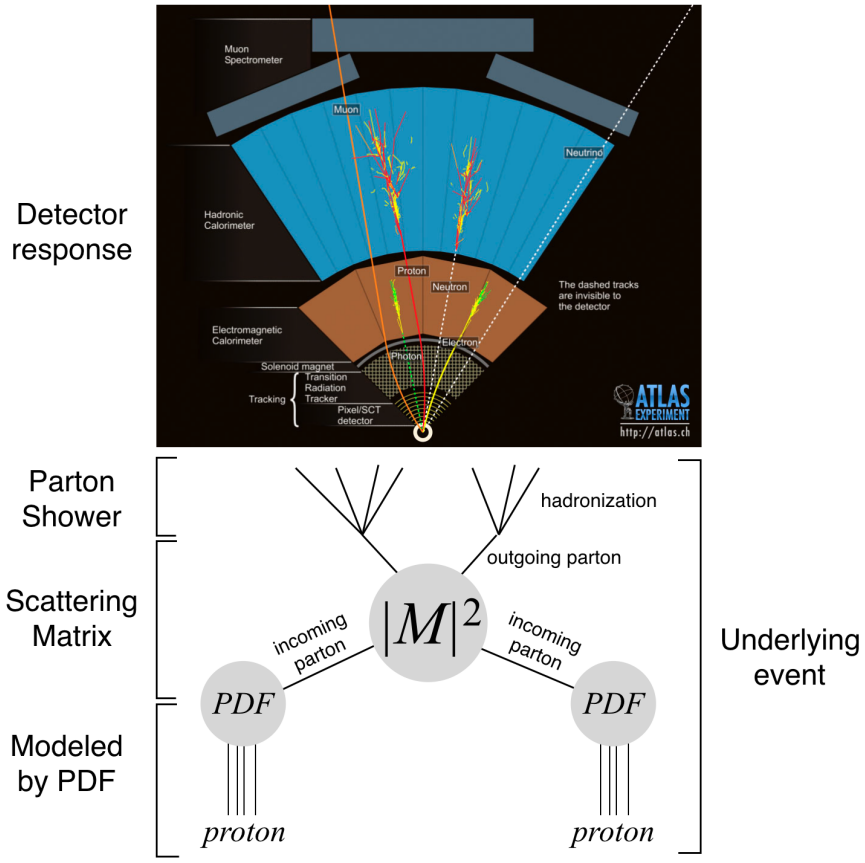


Figure 4.3: Schematic summary of simulating pp collision. The general strategy is that of *divide et impera*. The incoming partons from each protons are modeled by the parton distribution function (PDF). Several choices of PDFs are available in the market. The incoming partons to outgoing parton scattering processes are modeld by the scattering matrix $|\mathcal{M}|^2$ computed from Feynman rules of the Standard Model by various generators. The outgoing partons hadronizes to hadrons through parton shower programs The overall underlying events not directly accounted are modeled as well. The response of the detector to the particles hitting the detectors is modeled by GEANT4 or GEANT4 combined with a parametrized GEANT4-based calorimeter simulation [52]. (Top image: ATLAS Experiment © 2017 CERN).

hadronization of outgoing partons are modeled through parton shower programs. The overall underlying events not directly accounted are modeled as well. The response of the detector to the particles hitting the detectors is modeled by GEANT4 or GEANT4 combined with a parametrized GEANT4-based calorimeter simulation [52].

The scattering matrix level calculations are mostly performed by the next-to-leading order (NLO) generator POWHEG [54]. For processes, where high multiplicities of partons are pertinent, ALPGEN [55] or SHERPA [56] provide the tree level calculations up to several partons and the generated events are merged. For few cases, the ACERMC [57] and GG2VV [58] generators are used for small cross section process. The calculations for the scattering matrix level are matched to a parton shower program. Parton shower pro-

grams such as PYTHIA6 [59], PYTHIA8 [60], HERWIG [61] (with the underlying event modeled by JIMMY [62]), or SHERPA models not only the parton shower but also the hadronization and underlying events. The PDFs used are CT10 [63] for the POWHEG and SHERPA and CTEQ6L1 [64] for ALPGEN and ACERMC. The pile-up interactions are modeled with PYTHIA8 using the ATLAS A2 set of tuned parameters [65] and the MSTW2008LO PDF set [66]. The list of generators used and the cross section times the branching ratio of each process is shown in Table 4.1.

Table 4.1: The table and caption is taken from Ref. [18]. Monte Carlo samples used to model the signal and background processes. The corresponding cross sections times branching fractions, $\sigma \cdot \mathcal{B}$, are quoted at $\sqrt{s} = 8$ TeV. The branching fractions include the decays $t \to Wb$, $W \to \ell\nu$, and $Z \to \ell\ell$ (except for the process $ZZ \to \ell\ell\nu\nu$). Here ℓ refers to e, μ, or τ for signal and background processes. The neutral current $Z/\gamma^* \to \ell\ell$ process is denoted Z or γ^*, depending on the mass of the produced lepton pair. Vector-boson scattering (VBS) and vector-boson fusion (VBF) background processes include all leading-order diagrams with zero QCD vertices for the given final state (except for diagrams with Higgs bosons, which only appear in the signal processes).

Process		MC generator	$\sigma \cdot \mathcal{B}$ (pb)
Signal			
ggF	$H \to WW^*$	POWHEG+PYTHIA8	0.435
VBF	$H \to WW^*$	POWHEG+PYTHIA8	0.0356
VH	$H \to WW^*$	PYTHIA8	0.0253
WW			
$q\bar{q} \to WW$ and $qg \to WW$		POWHEG+PYTHIA6	5.68
$gg \to WW$		GG2VV+HERWIG	0.196
$(q\bar{q} \to W) + (q\bar{q} \to W)$		PYTHIA8	0.480
$q\bar{q} \to WW$		SHERPA	5.68
VBS $WW + 2$ jets		SHERPA	0.0397
Top quarks			
$t\bar{t}$		POWHEG+PYTHIA6	26.6
Wt		POWHEG+PYTHIA6	2.35
$tq\bar{b}$		ACERMC+PYTHIA6	28.4
$t\bar{b}$		POWHEG+PYTHIA6	1.82
Other dibosons (VV)			
$W\gamma$	($p_{\mathrm{T}}^{\gamma} > 8$ GeV)	ALPGEN+HERWIG	369
$W\gamma^*$	($m_{\ell\ell} \leq 7$ GeV)	SHERPA	12.2
WZ	($m_{\ell\ell} > 7$ GeV)	POWHEG+PYTHIA8	12.7
VBS $WZ + 2$ jets	($m_{\ell\ell} > 7$ GeV)	SHERPA	0.0126
$Z\gamma$	($p_{\mathrm{T}}^{\gamma} > 8$ GeV)	SHERPA	163
$Z\gamma^*$	(min. $m_{\ell\ell} \leq 4$ GeV)	SHERPA	7.31
ZZ	($m_{\ell\ell} > 4$ GeV)	POWHEG+PYTHIA8	0.733
$ZZ \to \ell\ell\nu\nu$	($m_{\ell\ell} > 4$ GeV)	POWHEG+PYTHIA8	0.504
Drell-Yan			
Z	($m_{\ell\ell} > 10$ GeV)	ALPGEN+HERWIG	16500
VBF $Z + 2$ jets	($m_{\ell\ell} > 7$ GeV)	SHERPA	5.36

Chapter 5

First Evidence for Vector Boson Fusion Produced $H \to WW^*$

The search for VBF produced $H \to WW^*$ is conducted using both the 2011 and 2012 pp collision data recorded by the ATLAS detector corresponding to an integrated luminosity of 25 fb^{-1} [18]. A boosted decision tree (BDT) method [67] (detailed in Sec. 5.2) is adopted for the first time in the search for VBF produced $H \to WW^* \to \ell\nu\ell\nu$. A cross-check of the BDT analysis with a more conventional selection-based analysis is conducted and has been shown to be consistent with the result of BDT based analysis.

The key feature to the analysis is the usage of BDT algorithm. The BDT algorithm combines several key characteristics of the VBF process and those of the Higgs decay into a single powerful discriminant O_{BDT}. The O_{BDT} variable is more effective than selecting events based on the individual variables inputted to the BDT. This in return provides a higher sensitivity than the conventional selection-based analysis. More details on the BDT analysis is documented in this chapter.

It is important to note that the search for ggF produced $H \to WW^*$ is performed in an orthogonal phase-space region in conjunction with the VBF analysis. The VBF analysis is specifically optimized for selecting VBF events, yet small but significant contributions from ggF process contaminates the VBF signal region; an accurate profile of the ggF contribution in VBF signal region is necessary. Orthogonal phase-space regions targeting ggF signal are constructed and from which the ggF contribution is estimated via extrapolation into VBF signal region.

A global likelihood fit to the data in key discriminant variable distributions including O_{BDT} is performed in both the VBF and ggF signal regions to obtain the signal strength of the VBF process. The signal strength is defined to be the ratio μ_{VBF}, which is the number of observed VBF process events to the number of expected VBF process. The μ_{VBF} is equal to one if the data agrees with the SM. The significance on excluding the hypothesis that no VBF process contribution is present (i.e. $\mu_{\mathrm{VBF}}/\mu_{\mathrm{ggF}} = 0$) in data is interpreted as the significance for the existence of VBF produced $H \to WW^*$ process. The ggF analysis is beyond the scope of this thesis and will not be discussed further.

5.1 Data Samples and Preselection

The final state of interest is two leptons ($\ell = e$ or μ) plus two or more jets. This analysis thus uses the events recorded by triggers requiring a single lepton or two leptons (dilepton) triggers. The details on trigger requirements for the 2012 data taking is as follows. Single lepton triggers required $p_{\mathrm{T}} \geq 24$ GeV regardless of the lepton flavor. The di-electron trigger required $p_{\mathrm{T}} \geq 12$ GeV on both electrons. The di-muon trigger required $p_{\mathrm{T}} \geq 18$ GeV on the leading muon and $p_{\mathrm{T}} \geq 8$ GeV on the subleading muon. The combined $e\mu$ trigger required $p_{\mathrm{T}} \geq 12$ GeV on the electron and $p_{\mathrm{T}} \geq 8$ GeV on the muon. Trigger strategy in 2011 data taking followed similar approach as the 2012 data taking. Majority of the differences in trigger strategy between 2011 and 2012 are in the p_{T} thresholds of trigger requirements. With lower average pile-up condition, the p_{T} threshold can be few GeV lower in 2011 data taking. Recorded data events are subjected to quality criteria, which require relevant detector components were operating correctly at the time of the event. After the data quality requirement, the integrated luminosity of pp collision events recorded resulted in 20.3 fb^{-1} at $\sqrt{s} = 8$ TeV for 2012 and 4.5 fb^{-1} at 7 TeV for 2011.

A collection of loose selection criteria are required on the events. They will be referred to as preselection throughout this chapter. The preselection criteria are as follows. Events with more than two leptons with $p_{\mathrm{T}} > 10$ GeV are rejected. The leading lepton (ℓ_1) is required to have $p_{\mathrm{T}} \geq 22$ GeV and the subleading lepton (ℓ_2) is required to have $p_{\mathrm{T}} \geq 10$ GeV. As the Higgs boson is a neutral particle, the two reconstructed leptons are required to have opposite charge. As the two leading jets for the signal events have distinct VBF-like topology, events are required to have $n_j \geq 2$. All jets are required to have $p_{\mathrm{T}}^j \geq 25$ GeV for $|\eta_j| < 2.4$ and $p_{\mathrm{T}}^j \geq 30$ GeV for $|\eta_j| \geq 2.4$.

Events with different-flavor leptons ($e\mu$) have significantly less Z/γ^* background, as the only contributions are from a small rate of fully leptonic decay of τ-lepton pairs with suppressed lepton momenta due to undetected neutrinos. Data samples with $e\mu$ events are thus separated out from the same-flavor ($ee/\mu\mu$) events, and the two samples are analyzed separately. Events are required to have an invariant mass $m_{\ell\ell} \geq$ 10 GeV (12 GeV) for $e\mu$ ($ee/\mu\mu$) samples in order to remove low-mass meson resonances. For $ee/\mu\mu$ events, an additional requirement of $m_{\ell\ell} < m_Z - 15$ GeV is applied to reduce events with resonant Z bosons. The prolific Z/γ^* events do not have neutrinos produced in the process. This results in low $E_{\mathrm{T}}^{\mathrm{miss}}$ or $p_{\mathrm{T}}^{\mathrm{miss}}$. Sample of $ee/\mu\mu$ events are thus required to have $E_{\mathrm{T}}^{\mathrm{miss}} > 45$ GeV and $p_{\mathrm{T}}^{\mathrm{miss}} > 40$ GeV. Sample of events after preselection with lepton flavor of $e\mu$ ($ee/\mu\mu$) is referred to as the $e\mu$ ($ee/\mu\mu$) sample throughout this chapter. This concludes the discussion of preselection requirements and the criteria are outlined in Table 5.1.

After preselection requirements are applied, the $e\mu$ and $ee/\mu\mu$ sample of 6.1×10^4 events and 2.7×10^4 events, respectively, are mainly consisted of top-quark backgrounds. The subdominant backgrounds are

Z/γ^* and WW. The breakdowns of each background process contribution for $e\mu$ and $ee/\mu\mu$ samples are presented in Table 5.4.

Table 5.1: Preselection criteria summary. The following selections are applied to all events considered in this analysis. Trigger and data quality requirements are not listed but also applied to all events. See text for more details on trigger and data quality requirements.

Objective	Selection
Object multiplicities and p_{T} thresholds	$p_{\mathrm{T}}^{\ell 1} > 22$ GeV and $p_{\mathrm{T}}^{\ell 2} > 10$ GeV Veto events with $n_\ell > 2$ with $p_{\mathrm{T}}^{\ell 3} > 10$ GeV Opposite-charge leptons $n_j \geq 2$ $p_{\mathrm{T}}^{j} \geq 25$ GeV for $\lvert\eta_j\rvert < 2.4$ $p_{\mathrm{T}}^{j} \geq 30$ GeV for $\lvert\eta_j\rvert \geq 2.4$
reject Z/γ^* background ($ee/\mu\mu$ events only)	$m_{\ell\ell} < m_Z - 15$ GeV $E_{\mathrm{T}}^{\mathrm{miss}} > 45$ GeV $p_{\mathrm{T}}^{\mathrm{miss}} > 40$ GeV
reject low-mass meson resonances	$m_{\ell\ell} > 10$ GeV (12 GeV) for $e\mu$ ($ee/\mu\mu$) events

5.2 Boosted Decision Tree

A boosted decision tree algorithm can classify events or objects based on several discriminating input variables. During the training phase, a *forest* of binary decision trees that classify a given event as signal or background are grown through an iterative process. Taking a weighted average of all the trees result in a single BDT output variable O_{BDT}, which classifies events more accurately than any single tree in the forest would. This process of combining multiple weak classifiers into a single powerful classifier is called the *boosting* [67].

5.2.1 Algorithm

Each binary decision tree is grown from a root node, which contains all events. The algorithm first searches the best input variable x that maximally separates signal and background events in a given node. Each node is then split into two sub-nodes by events that satisfy $x > a$ or $x \leq a$. A node stops splitting if one of the stop criteria is met; this node is called a *leaf* as it is at an end of a tree branch. Each leaf is labeled as a signal or background leaf depending on its purity of signal events in the leaf. Once all nodes stop splitting a single binary decision tree is fully grown. A binary decision tree can be represented as a boolean function h, where it returns $+1$ for events in signal leaves or -1 for events in background leaves.

It is inevitable that a subset of signal events fall under background leaves and vice versa. Between each iteration of growing a *binary tree*, incorrectly assigned signal (background) events are weighted higher (lower); this process is called the *boost*. The next iteration grows a tree that adopts the incorrectly assigned events more accurately due to their boosted weights. This process is repeated until a desired number of trees are grown.

There are few tunable parameters in the training phase; the tree depth parameter limits the maximum depth of each tree; each node can be required to have minimum number of events to ensure that each nodes are statistically significant; the threshold on the purity requirement in labeling a terminating node a signal or background node can be tuned. The optimal parameter choice for the VBF analysis is obtained through a brute force scan over possible parameter configurations.

To classify data events as signal or background events from a forest of binary trees, the weighted average of each binary tree function $h_i(\boldsymbol{x})$ is taken as follows [68],

$$O_{\mathrm{BDT}}(\boldsymbol{x};\boldsymbol{w}) = \sum_{i=0}^{N} w_i h_i(\boldsymbol{x}), \tag{5.1}$$

where w_i is the weight per binary tree function $h_i(\boldsymbol{x})$, $\boldsymbol{w} = (w_0, w_1, ..., w_N)$ is an array of weight values, N is the total number of trees in a forest, and $\boldsymbol{x} = (x_1, x_2, ..., x_M)$ is an array of M input variable values for a given event. The weights w_i are obtained during the training phase by minimizing the loss-function $L = \ln\left(1 + e^{-2\hat{O}_{\mathrm{BDT}}(\boldsymbol{x})y}\right)$ in each iteration, where y is the true value of classifier output (i.e. $+1$ for signal, -1 for background) and the $\hat{O}_{\mathrm{BDT}}$ is the trained classifier output up to the present iteration. The TMVA package [68] was used to perform the training and compute O_{BDT} using trained knowledge.

5.2.2 Input Variables

The kinematics of leptons are governed by the dynamics of Higgs boson decay. Fig. 5.1 illustrates the kinematics of the $H \to WW^* \to \ell\nu\ell\nu$ decay. As the spin-0 Higgs boson decays into two spin-1 W bosons, the spin angular momentum conservation dictates that directions of spins of W bosons to be opposite to each other. This is illustrated by the large double arrows pointing in opposite directions in Fig. 5.1. Since only left-handed neutrinos (right-handed anti-neutrinos) couple to W^+ (W^-) boson, directions of spins of the leptons will tend to be aligned; this leads to a pair of charged leptons with a small opening angle [69, 70]. Figs. 5.2(a) and 5.2(b) shows the distribution of azimuthal opening angle $\Delta\phi_{\ell\ell}$, between the two leptons for $e\mu$ events and $ee/\mu\mu$ events after event preselection, respectively. The signal process shown in solid black line peaks at low $\Delta\phi_{\ell\ell}$, indicating small opening angle between the two leptons. As both leptons come from the

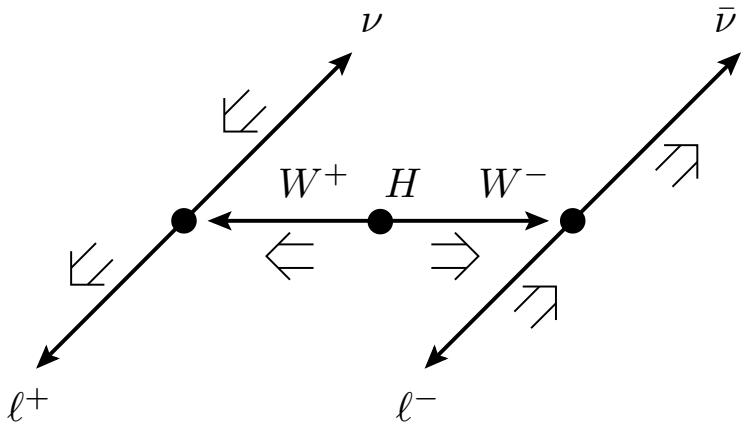


Figure 5.1: Illustration of $H \to WW^*$ decay from Ref. [18]. The solid single arrows indicate the particles directions of motion and the large double arrows indicate their spin projections. Due to the spin angular momentum conservation, the W bosons have opposite spins, while leptons from each W boson have aligned spins. The Higgs and W boson decays are shown in the decaying particles rest frame. Because of the chiral nature of W boson decay, the charged leptons have a small opening angle in the laboratory frame.

decay of a resonant Higgs boson, the invariant mass distribution of the lepton pair tends to be less than the half of the Higgs mass. This can be seen in Figs. 5.2(c) and 5.2(d), which show the distribution of $m_{\ell\ell}$ in $e\mu$ events and $ee/\mu\mu$ events, respectively. This feature is useful for suppressing background processes with high momentum transfer Q^2 that tends to have a long tail in $m_{\ell\ell}$ distribution, such as pair-produced top-quark process. Because of the two neutrinos produced in the decay of the Higgs boson, a full reconstruction of the invariant mass of the Higgs boson is not possible. Instead a "transverse mass" m_{T} [71] is computed:

$$m_{\mathrm{T}} = \sqrt{\left(E_{\mathrm{T}}^{\ell\ell} + p_{\mathrm{T}}^{\nu\nu}\right)^2 - \left|\boldsymbol{p}_{\mathrm{T}}^{\ell\ell} + \boldsymbol{p}_{\mathrm{T}}^{\nu\nu}\right|^2}, \tag{5.2}$$

where $E_{\mathrm{T}}^{\ell\ell} = \sqrt{(p_{\mathrm{T}}^{\ell\ell})^2 + (m_{\ell\ell})^2}$, $\boldsymbol{p}_{\mathrm{T}}^{\ell\ell}$ ($\boldsymbol{p}_{\mathrm{T}}^{\ell\ell}$) is the vector sum of the neutrino (lepton) transverse momenta, and $p_{\mathrm{T}}^{\nu\nu}$ ($p_{\mathrm{T}}^{\ell\ell}$) is its modulus [18]. The m_{T} distribution exhibits a peak structure just below the Higgs boson mass. More importantly a clear upper bound at the Higgs boson mass as shown in Figs. 5.2(e) and 5.2(f). The three variables, $\Delta\phi_{\ell\ell}$, $m_{\ell\ell}$, and m_{T} are provided as inputs to the BDT algorithm.

The two leading jets in the events are referred to as VBF tagged jets throughout this chapter. The kinematics of VBF tagged jets are very distinct. The two leading jets for the VBF process tend to have large rapidity gap $\Delta y_{jj} = |y_{j1} - y_{j2}|$. This can be seen in Figs. 5.3(a) and 5.3(b). The VBF jets are also highly energetic and this leads to a large tail in invariant mass m_{jj} distribution. This can also be seen in Figs. 5.3(c) and 5.3(d). The Δy_{jj} and m_{jj} variables are also provided as inputs to the BDT algorithm. The rapidity gap between the two VBF tagged jets is referred to as the central region. The decay products from the VBF produced Higgs boson tends to populate the central region. The centrality of a given lepton is

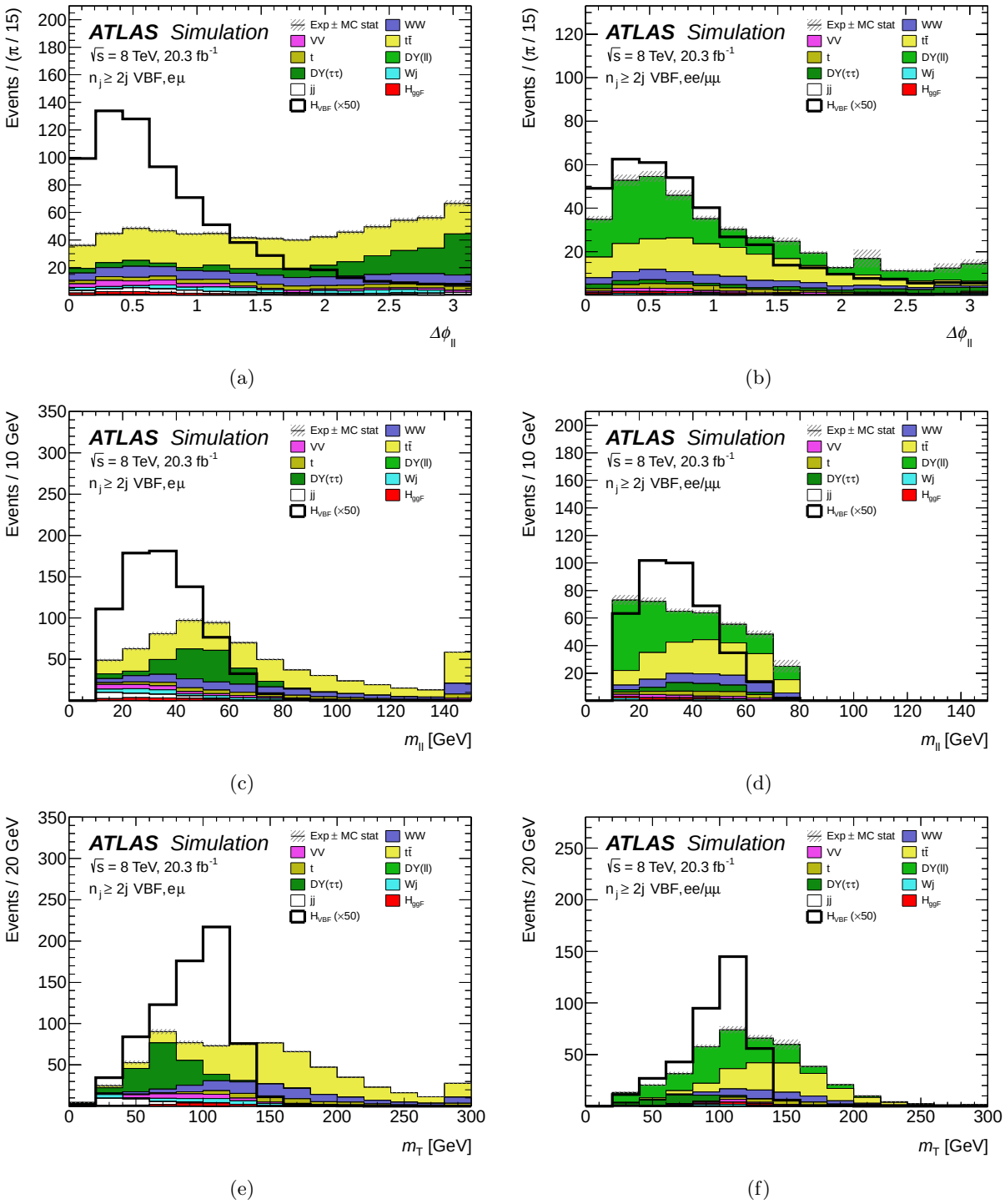


Figure 5.2: Key kinematic variables discriminating $H \to WW^* \to \ell\nu\ell\nu$ against background events by exploting characteristics of the Higgs boson decay. The left columns show $e\mu$ sample events, and the right $ee/\mu\mu$ sample events. The top row plots show the azimuthal angle difference between the two leptons. The middle row plots show the invariant mass of the two lepton system. The bottom row plots show the transverse mass of the two lepton plus missing transverse momentum system.

defined as,

$$C_\ell = \left| \eta_\ell - \frac{\sum \eta_{jj}}{2} \right| \Big/ \Delta\eta_{jj}, \tag{5.3}$$

where η_ℓ is the given lepton's pseudorapidity and $\sum \eta_{jj} = \eta_{j1} + \eta_{j2}$. When $C_\ell = 1$, the lepton is pointing along one of the VBF tagged jet. The sum of the two lepton centrality $\sum C_\ell = C_{\ell 1} + C_{\ell 2}$ is provided as an input to the BDT algorithm. The distribution of $\sum C_\ell$ can be seen in Figs. 5.3(e) and 5.3(f). The centralness of decay products with respect to the jets implies that the invariant mass between a lepton and a VBF tagged jet will be large. The sum of all pairs between leptons and VBF tagged jets $\Sigma\, m_{\ell j} = m_{\ell 1,j1} + m_{\ell 1,j2} + m_{\ell 2,j1} + m_{\ell 2,j2}$ is provided as an input to the BDT algorithm. Figs. 5.4(a) and 5.4(b) shows the $\Sigma\, m_{\ell j}$ distribution. The distribution of $\Sigma\, m_{\ell j}$ for VBF signal process on average has larger values than for background processes. The momentum balance $p_{\mathrm{T}}^{\mathrm{sum}} = |\boldsymbol{p}_{\mathrm{T}}^{\mathrm{sum}}| = |\boldsymbol{p}_{\mathrm{T}}^{\ell\ell} + \boldsymbol{p}_{\mathrm{T}}^{\mathrm{miss}} + \sum \boldsymbol{p}_{\mathrm{T}}^{j}|$ is a variable sensitive to extra low p_{T} jet activity. The jets that are summed over in $\sum p_{\mathrm{T}}^{j}$ are required to have $p_{\mathrm{T}} \geq 20$ GeV. Jets with $p_{\mathrm{T}} < 20$ GeV in an event is not included in the calculation of $p_{\mathrm{T}}^{\mathrm{sum}}$ leading to larger momentum imbalance. Jets in the VBF signal process are mostly energetic while background processes can have significant low p_{T} jets from QCD radiations. Figs. 5.4(c) and 5.4(d) shows on average lower momentum imbalance compared to other background processes.

5.2.3 Boosted Decision Tree Discriminant

Fig. 5.5 shows O_{BDT} distribution after applying the preselection (Table 5.1) and the event selection (Table 5.2) that is discussed in the next section. The VBF signal process contribution is shown in solid black line. The background contributions are stacked. The ggF Higgs process contribution is treated as background. The VBF signal contribution is overlaid to show the differences in shapes compared to the background contributions. The output distribution shows a good separation between signal and background processes.

5.3 Event Selection

In addition to the preselection, few more selections are applied for further suppression of background. Events are required to have $n_b = 0$ to reduce the dominant top-quark background. The invariant mass of fully leptonically decaying di-tau system from $Z/\gamma^* \to \tau\tau$ process is difficult to reconstruct due to missed neutral particles. With the approximation that the neutral particles from the tau decay are largely collinear with the charged lepton decay product, the di-tau invariant mass $m_{\tau\tau}$ can be calculated [72–74]. To suppress $Z/\gamma^* \to \tau\tau$ process, events with $m_{\tau\tau} \geq m_Z - 25$ GeV are removed*.

*This requirement also maintains orthogonality with fully leptonically decaying $H \to \tau\tau$ analysis for global Higgs coupling combination.

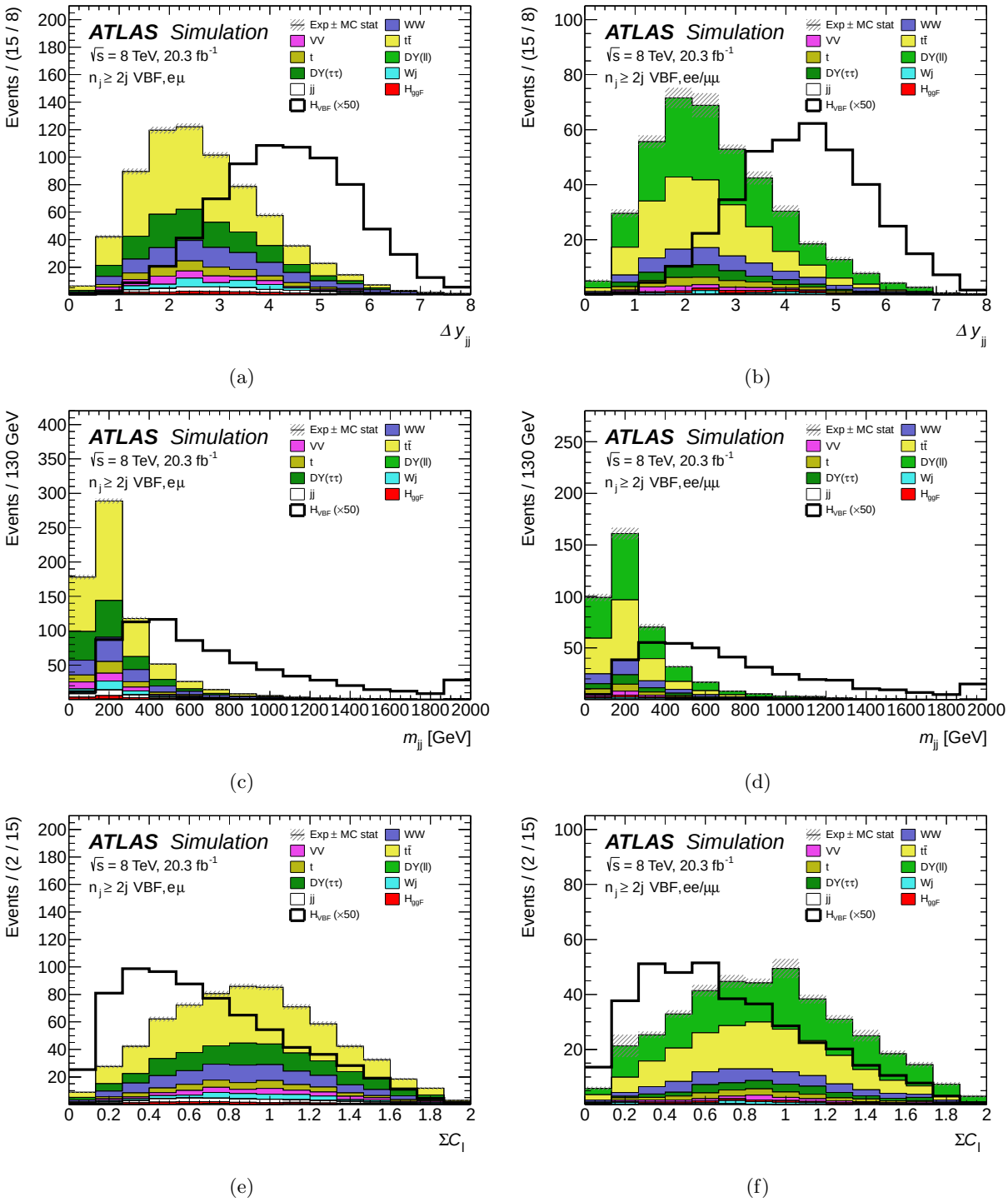


Figure 5.3: Key kinematic variables discriminating VBF process against background events by exploting characteristics of VBF process. The left columns show $e\mu$ sample events, and the right $ee/\mu\mu$ sample events. The top row plots show the rapidity gap between the two VBF tagged jets. The middle row plots show the invariant mass of the two VBF tagged jets system. The bottom row plots show the sum of centralities of the leptons. See text for details on the definition of centrality.

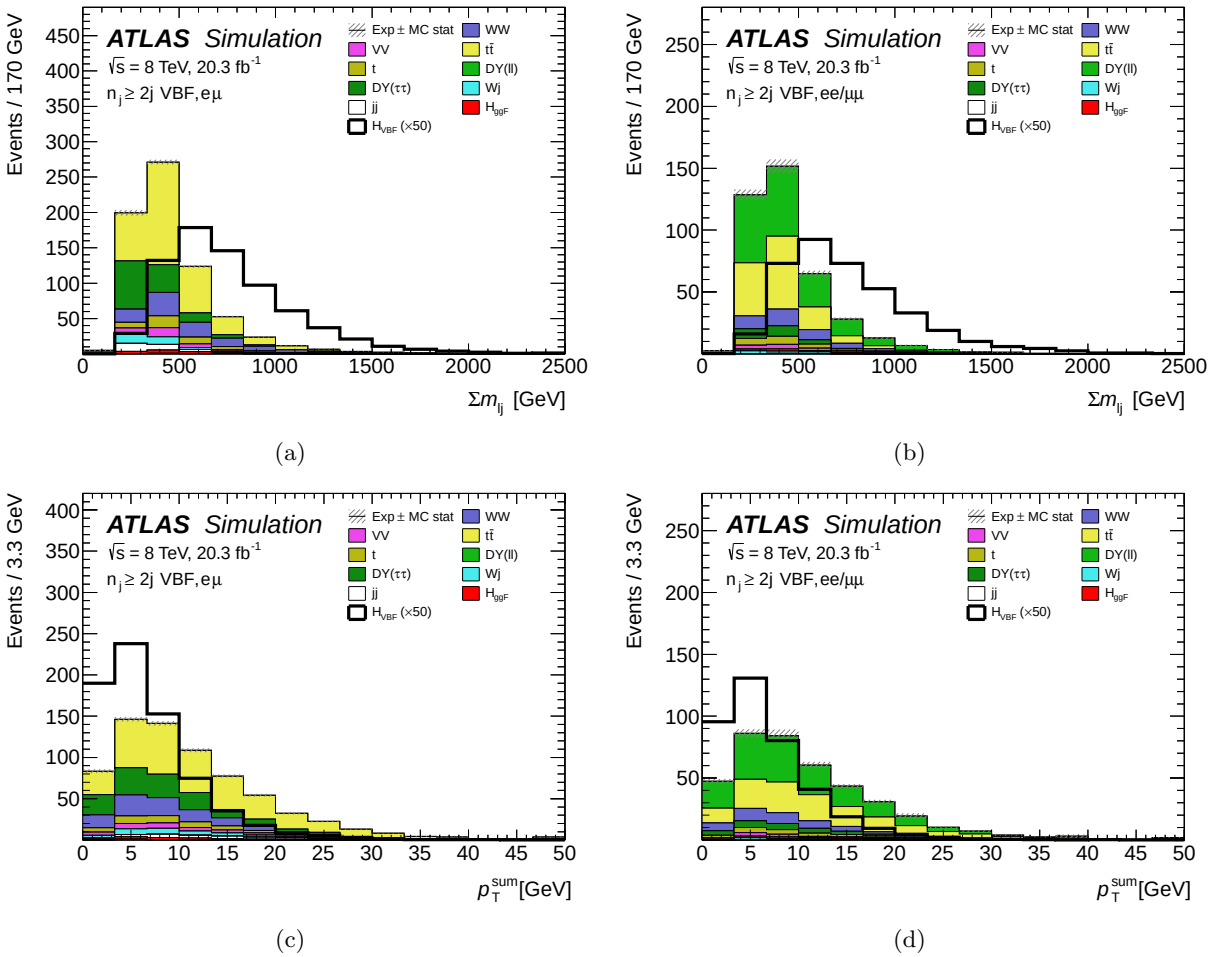


Figure 5.4: Key kinematic variables discriminating VBF $H \to WW^* \to \ell\nu\ell\nu$ process against background events by exploiting characteristics of the VBF $H \to WW^* \to \ell\nu\ell\nu$ process. The left columns show $e\mu$ sample events, and the right $ee/\mu\mu$ sample events. The top row plots show the sum of invariant masses of lepton-jet pairs from two leptons and two VBF tagged jets. The bottom row plots show the momentum imbalance of the system. See text for details on the definition of the momentum imbalance.

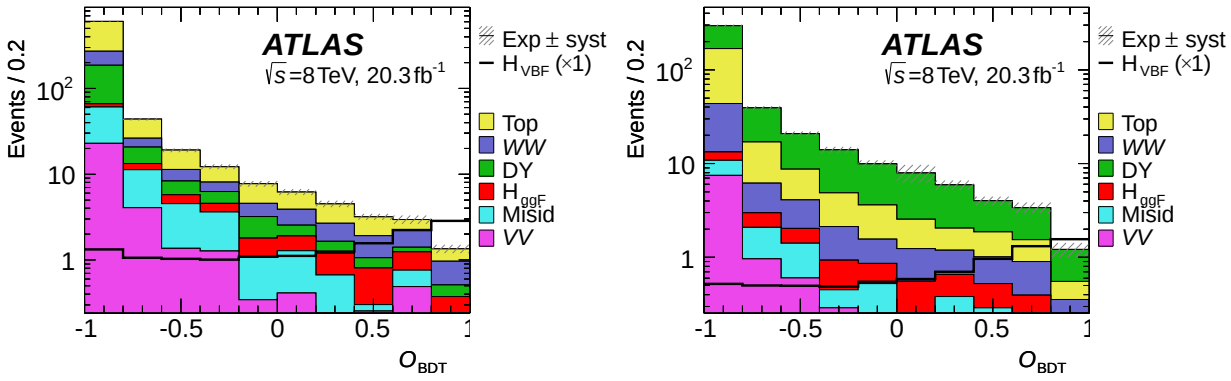


Figure 5.5: Output O_{BDT} distribution for (a) $e\mu$ and (b) $ee/\mu\mu$ samples. Background shows a steeply falling distribution as O_{BDT} approaches 1, while the VBF signal process shows a rise.

In the central region defined by the VBF tagged jets, a low hadronic activity in the central region is expected for the VBF signal process as the mediating vector bosons do not carry color charge. Events are thus required to not have extra jets in the central region. This is called the "central jet veto" [75]. The central jet veto uses jets with $p_{\mathrm{T}} > 20$ GeV. The veto is effective in reducing background processes with jets produced via QCD radiation. Events are required to have $C_{\ell 1} < 1$ and $C_{\ell 2} < 1$[†]. Table 5.2 summarizes the event selection.

Table 5.2: Event selection criteria summary.

Objective	Selection
reject top-quark and $Z/\gamma^* \to \tau\tau$ background	$n_b = 0$ $m_{\tau\tau} < m_Z - 25$ GeV
VBF characteristics in the central region	Central jet veto $C_{\ell 1} < 1$ and $C_{\ell 2} < 1$

All selections relevant for the nominal BDT analysis have been discussed. Table 5.3(a) shows the yields in $e\mu$ and $ee/\mu\mu$ samples after the preselection (Table 5.1) and the BDT event selection (Table 5.2). The expected number of signal events is expected to be about 2% of the total expected background events. The O_{BDT} distribution is binned into four bins. The bin boundaries were chosen to maximize significance while keeping each bin sufficiently populated. The final bin boundary configuration chosen is $[-1, -0.48, 0.3, 0.78, 1]$, while each bin is labeled a number starting from 0 to 3, respectively. Events in BDT bin 0 kinematically do not resemble VBF topology and are discarded[‡] in the final fit. The BDT bin 1–3 show increasing signal-to-background ratio. The expected number of signal events N_{VBF} in the BDT bin 3 of $e\mu$ is 3.1 while the expected number of background events N_{bkg} is 1.5. The $N_{\mathrm{VBF}}/N_{\mathrm{bkg}}$ is approximately 2, indicating a very pure signal phase-space region. It is important to note that N_{bkg} includes the expected number of ggF produced Higgs boson events N_{ggF}, as the analysis considers it a background source. The conclusion in $ee/\mu\mu$ sample is similar with signal yield roughly halved. The difference is due to the additional requirement on the missing energy to suppress large Z/γ^* background that results in lower acceptance of signal and background events.

While the BDT analysis makes use of the O_{BDT} that combines key discriminating variables, the cross-check analysis makes explicit selections to the events using the BDT input variables. The same preselection and event selection for the BDT analysis is applied to the cross-check analysis. In addition, events in the cross-check analysis are required to have $p_{\mathrm{T}}^{\mathrm{sum}} < 15$ GeV, $m_{jj} > 600$ GeV, $\Delta y_{jj} > 3.6$, $\Delta\phi_{\ell\ell} < 1.8$,

[†]This is also known as "outside lepton veto".
[‡]Events in BDT bin 0 are, however, used in the ggF analysis.

Table 5.3: Summary table of number of expected and observed events in the BDT analysis reproduced from Ref. [18]. The yield N_{bkg} includes N_{ggF} as ggF process is treated as a background. (a) Table of expected yields in each contribution after the preselection (Table 5.1) and the BDT event selection (Table 5.2) except $O_{\text{BDT}} > -0.48$ applied. (b) Table of the yields in each BDT bin. Only the BDT bins 1–3 are used in the fit.

	Summary					Composition of N_{bkg}						
Selection	$N_{\text{obs}}/N_{\text{bkg}}$	N_{obs}	N_{bkg}	N_{ggF}	N_{VBF}	N_{WW}	N_{top}	N_{misid}	N_{VV}	$N_{ee/\mu\mu}$	$N_{\tau\tau}^{\text{QCD}}$	$N_{\tau\tau}^{\text{EW}}$
(a) Before the BDT classification												
$e\mu$ sample	1.04 ± 0.04	718	689	15	15	101	369	52	31	2.2	130	2
$ee/\mu\mu$ sample	1.18 ± 0.08	469	397	6.9	7.7	40	149	6.4	10.1	168	23	1
(b) Bins in O_{BDT}												
$e\mu$ sample												
Bin 0 (not used)	1.02 ± 0.04	661	650	8.8	3.0	92	352	47	28	2.2	126	1
Bin 1	0.91 ± 0.15	37	41	3.0	4.2	6.0	19.7	5.1	2.6	-	3.9	0.2
Bin 2	1.89 ± 0.52	14	7.4	1.2	4.2	2.1	2.1	0.7	0.8	-	0.3	0.4
Bin 3	4.12 ± 1.73	6	1.5	0.4	3.1	0.5	0.3	-	0.1	-	0.1	0.1
$ee/\mu\mu$ sample												
Bin 0 (not used)	1.91 ± 0.08	396	345	3.8	1.3	35	139	5.2	8.8	137	20.5	0.5
Bin 1	1.13 ± 0.17	53	47	1.5	2.2	3.5	12.2	1.0	0.9	26	1.7	0.1
Bin 2	1.65 ± 0.46	14	8.5	0.6	2.5	1.1	1.3	0.2	0.3	4.4	0.3	0.1
Bin 3	5.56 ± 2.39	6	1.1	0.2	1.7	0.3	0.2	-	-	0.4	-	0.1

$m_{\ell\ell} < 50$ GeV, and $m_{\text{T}} < 150$ GeV. The thresholds on the selections were optimized to have maximum expected significance to the VBF signal. Table 5.4 shows the expected yields of each component for the cross-check analysis. The table is broken into two for $e\mu$ sample and $ee/\mu\mu$ sample. As more selections are applied the signal-to-background ratio improves and the ratio $N_{\text{obs}}/N_{\text{bkg}}$ deviates from unity indicating a presence of signal process in data events. With all requirements are applied, the expected number of signal events for $e\mu$ sample is 4.7, while the expected number of total background is 5.5. The expected number of signal events for $ee/\mu\mu$ sample is 2.2, while the expected number of total background is 3.7. The expected number of signal and background events for $ee/\mu\mu$ and the signal-to-background ratio are relatively smaller than those of $e\mu$ sample for the same reason as the BDT analysis that a more stringent missing energy requirement is applied to the $ee/\mu\mu$ sample.

5.4 Background Estimation and Their Uncertainties

All MC based models of the SM are an approximation of the true calculation of the scattering process. Some disagreement between the data distribution and the predicted distribution by the MC generators are expected. Whenever a clear arguments can be made, using the data distribution in an orthogonal control region (CR) to predict expected background distribution in the signal region (SR) is preferred. Table 5.3

Table 5.4: Summary table of number of expected and observed events in the cross-check analysis reproduced from Ref. [18]. The yield N_{bkg} includes N_{ggF} as ggF process is treated as a background. The breakdown of each contribution is shown for $e\mu$ sample first and then for $ee/\mu\mu$ sample. The expected yields for the row, $e\mu$ ($ee/\mu\mu$) sample, is after the preselection requirement. The following rows show the breakdown of each contribution after applying the selection and all previous selections listed in addition to the preselection. The selections listed are for the cross-check analysis. For the nominal BDT analysis, several discriminating variables used to make selection criteria for the cross-check analysis are used as inputs to the BDT algorithm instead.

	Summary					Composition of N_{bkg}						
Selection	$N_{\text{obs}}/N_{\text{bkg}}$	N_{obs}	N_{bkg}	N_{ggF}	N_{VBF}	N_{WW}	N_{top}	N_{misid}	N_{VV}	$N_{ee/\mu\mu}$	$N_{\tau\tau}^{\text{QCD}}$	$N_{\tau\tau}^{\text{EW}}$
$e\mu$ sample	1.00 ± 0.00	61434	61180	85	32	1418	54780	1155	380	51	3260	46
$n_b = 0$	1.02 ± 0.01	7818	7700	63	26	1036	3367	506	273	35	2400	29
$p_{\text{T}}^{\text{sum}} < 15$	1.03 ± 0.01	5787	5630	46	23	819	2180	323	201	27	2010	23
$m_{\tau\tau} < m_Z - 25$	1.05 ± 0.02	3129	2970	40	20	506	1447	207	132	7.6	627	5.8
$m_{jj} > 600$	1.31 ± 0.12	131	100	2.3	8.2	26.9	45.3	4.2	5.1	0.1	15	1.0
$\Delta y_{jj} > 3.6$	1.33 ± 0.13	107	80	2.1	7.9	18.6	40.0	3.9	3.3	-	11.6	0.8
$C_{j3} > 1$	1.36 ± 0.18	58	43	1.3	6.6	12.5	17.0	2.6	2.0	-	6.8	0.6
$C_{\ell 1} < 1$, $C_{\ell 2} < 1$	1.42 ± 0.20	51	36	1.2	6.4	11.1	13.3	2.6	1.6	-	5.7	0.6
$m_{\ell\ell}, \Delta\phi_{\ell\ell}, m_{\text{T}}$	2.53 ± 0.71	14	5.5	0.8	4.7	1.5	1.4	0.6	0.6	-	0.5	0.2
$ee/\mu\mu$ sample	0.99 ± 0.01	26949	27190	31	14	631	24760	239	137	690	679	16
$n_b, p_{\text{T}}^{\text{sum}}, m_{\tau\tau}$	1.03 ± 0.03	1344	1310	13	8.0	241	719	27	45	187	76	1.5
$m_{jj}, \Delta y_{jj}, C_{j3}, C_\ell$	1.39 ± 0.28	26	19	0.4	2.9	6.2	6.5	0.2	0.7	3.8	0.7	0.1
$m_{\ell\ell}, \Delta\phi_{\ell\ell}, m_{\text{T}}$	1.63 ± 0.69	6	3.7	0.3	2.2	0.6	0.8	0.2	0.1	1.5	0.3	0.1

shows that the dominant background sources in the most sensitive BDT bin (i.e. the BDT bin 3 in $e\mu$ sample) are WW, ggF Higgs boson, top-quark, and Z/γ^* backgrounds. This section discusses the estimation of these backgrounds and their uncertainties.

5.4.1 Summary of Background Estimation Methods

The goal of the background estimation for a given background is to obtain the expected number of the background in i-th bin of the kinematic variable X, represented as $\hat{B}_{\text{SR}}^{X_i}$. The number of background events in the MC sample (data) in the region $R \in \{\text{SR}, \text{CR}, \text{SR}', \text{CR}'\}$ and in the i-th bin of the kinematic variable X is represented as $B_R^{X_i}$ ($N_R^{X_i}$). The hat notation is used to distinguish between the final estimated number of backgrounds and the bare MC estimation. For the rest of the section the superscript X_i will be kept implicit unless explicitly stated. All the background estimation methods in the analysis follow one of the

following formulas,

$$\hat{B}_{\mathrm{SR}} = \underbrace{B_{\mathrm{SR}}}_{\text{Estimated directly by MC}}, \qquad \text{(Method A)} \tag{5.4}$$

$$= \underbrace{B_{\mathrm{SR}}}_{\text{Shape from MC}} \cdot \underbrace{\bar{N}_{\mathrm{CR}}/\bar{B}_{\mathrm{CR}}}_{\text{Normalized by data}}, \qquad \text{(Method B)} \tag{5.5}$$

$$= \underbrace{N_{\mathrm{CR}}}_{\text{Modeling from data}} \cdot \underbrace{B_{\mathrm{SR}}/B_{\mathrm{CR}}}_{\text{Extrapolated by MC}}, \qquad \text{(Method C)} \tag{5.6}$$

$$= \underbrace{N_{\mathrm{CR}}}_{\text{Modeling from data}} \cdot \underbrace{N_{\mathrm{SR}'}/N_{\mathrm{CR}'}}_{\text{Extrapolated by data}} \cdot \underbrace{\frac{B_{\mathrm{SR}}/B_{\mathrm{CR}}}{B_{\mathrm{SR}'}/B_{\mathrm{CR}'}}}_{\text{Correction by MC}}, \qquad \text{(Method D)} \tag{5.7}$$

where $\bar{N}_{\mathrm{CR}} = \sum_j N_{\mathrm{CR}}^{X_j}$ and $\bar{B}_{\mathrm{CR}} = \sum_j B_{\mathrm{CR}}^{X_j}$. Method A uses the MC generator directly to estimate the expected number of background. Method B uses the MC sample to model kinematic variables but the overall normalization is obtained from a separate control region by comparing the total number of observed background events $\bar{N}_{\mathrm{CR}}$ to the total number of expected background events from the MC sample $\bar{B}_{\mathrm{CR}}$ in the control region. Method C lets the observed background data events in CR to model a given kinematic distribution in the signal region by extrapolating from the control region to the signal region bin-by-bin by a factor α. The extrapolation factor for each bin is the ratio $\alpha = B_{\mathrm{SR}}/B_{\mathrm{CR}}$, where the numerator and the denominator are obtained from MC. Method D is an extension of Method C, where the extrapolation factor $\alpha = N_{\mathrm{SR}'}/N_{\mathrm{CR}'}$ is obtained from data in a separate region SR′ (CR′) that resembles the SR (CR). Method D assumes that the extrapolation factor obtained from the separate regions $N_{\mathrm{SR}'}$ and $N_{\mathrm{CR}'}$ are the same as the true extrapolation factor. (i.e. $N_{\mathrm{SR}'}/N_{\mathrm{CR}'} = N_{\mathrm{SR}}/N_{\mathrm{CR}}$, where N_{SR} is the true number of background events in SR.) The assumption can be checked using MC samples by verifying $f \equiv \frac{B_{\mathrm{SR}}/B_{\mathrm{CR}}}{B_{\mathrm{SR}'}/B_{\mathrm{CR}'}} = 1$ within a reasonable accuracy. If $f \neq 1$, the difference can be corrected by applying the correction factor f to the estimation. As MC calculations are approximations to the true calculations, theoretical uncertainties from MC calculations propagate to the background estimation uncertainty, wherever B_R shows up in the equation. An ideal background estimation minimizes the total uncertainty on a background by canceling out the theoretical uncertainty from the ratio $\alpha \equiv B_{\mathrm{SR}}/B_{\mathrm{CR}}$ while also minimizing statistical uncertainty on each N_R's.

Table 5.6 summarizes the background estimation used for each background sources and the dominant sources of uncertainty on each background. In addition to the background listed in the Table 5.6, there are diboson backgrounds (e.g. WZ, ZZ, etc.) and fake backgrounds (e.g. QCD and W). Estimation on number of expected diboson backgrounds is directly taken from the corresponding MC samples (Method A).

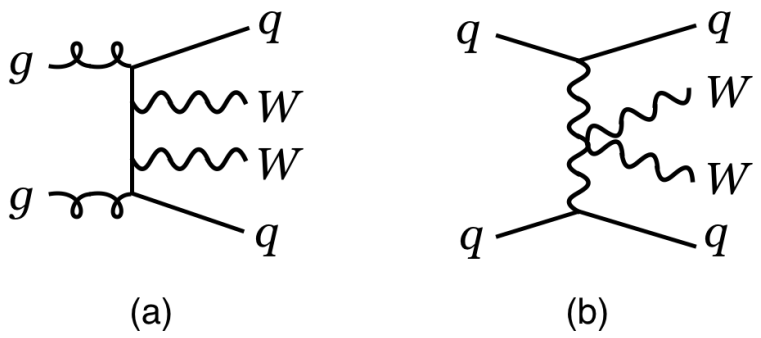


Figure 5.6: Feynman diagram examples for WW + 2 jets background with (a) QCD vertices and (b) electroweak vertices. The cross section for a class of diagrams similar to shown in (a) are at least on order of magnitude larger than that of (b).

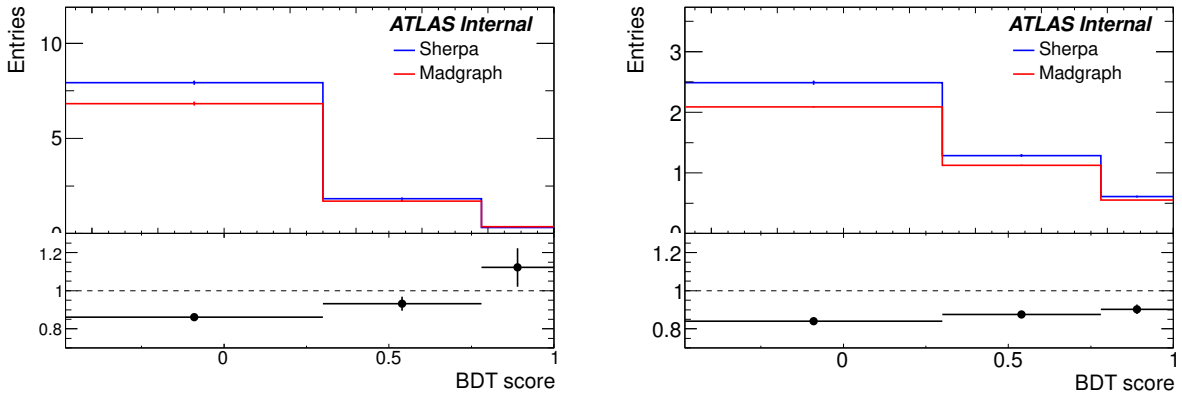
Fake backgrounds are estimated using a fake-factor method [18]. The method measures the misidentification rate of jets as leptons and apply the rate to a control sample to extrapolate and estimate the number of expected yields for QCD or W background in the signal region (Method D). The diboson backgrounds and fake backgrounds are subdominant in the VBF analysis and are not discussed in detail.

Table 5.5: Summary of dominant background sources. The dominant source of uncertainty on each background is also shown.

Background	Estimation Method	Dominant source of uncertainty
WW	Method A	QCD scale and modeling
ggF produced Higgs boson	Method B	QCD scale and selection on n_j
top-quark	Method C	Extrapolation from CR
$Z/\gamma^* \to \tau\tau$	Method B	Data statistics in control region
$Z/\gamma^* \to ee$, $\mu\mu$ ($ee/\mu\mu$ sample only)	Method D	Data statistics in control region

5.4.2 WW Background

There WW background in the VBF analysis enters into the signal region with two additional jets. The two additional jets can be generated from QCD or electroweak vertices. Fig. 5.6 shows an example Feynman diagram for each case. The jets from electroweak vertices tend to take more similar characteristics as the VBF tagged jets and thus enter into the VBF signal region with a relatively higher efficiency; although the cross section of the WW background with QCD vertices are an order of magnitude larger, around an equal number of contributions from each is in the most sensitive BDT bin. This can be understood by noting the resemblance of the Feynman diagram shown in Fig. 5.6 with that of VBF signal.



(a) (b)

Figure 5.7: Comparison on the O_{BDT} distribution for MC events generated by SHERPA and MADGRAPH for WW background with (a) QCD vertices and (b) electroweak vertices. Each set of events from the generators are scaled to the appropriate cross-section and integrated luminosity. The deviation from unity in each bin is taken as the systematic uncertainty [76].

The WW background is estimated by Method A using the SHERPA generator. The QCD and electroweak contributions are separately estimated. Comparison between SHERPA and MADGRAPH generated WW process with QCD vertices are shown in Fig. 5.7(a). Parton level kinematics were used in calculating the O_{BDT} distribution. The binning choices are kept same as the analysis. Difference between predictions from SHERPA and MADGRAPH of $8\% - 14\%$ is observed and taken as a systematic uncertainty. The renormalization and factorization scales are varied for MADGRAPH to estimate higher order contributions and the difference of 27% is taken as a systematic uncertainty. A similar comparison for WW process with electroweak vertices are shown in Fig. 5.7(b) and the difference of $10\% - 16\%$ is taken as a systematic uncertainty.

The MC prediction is validated using a kinematic selection that provides a reasonably pure sample of WW events. Events in $e\mu$ sample are required to pass preselection criteria (Table 5.1) and $n_b = 0$. An additional requirement of $m_{\mathrm{T}} > 100$ GeV is applied to enhance the WW contribution. To suppress top-quark background, events are required to have $m_{\mathrm{T2}} > 160$ GeV. The m_{T2} variable is an estimator of the mass of pair produced particles with neutral particles in the final states [77]. The m_{T2} can estimate the top-quark mass for fully leptonically decaying top-quark pair background. For a given $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution, possible neutrino combination is scanned to calculate m_{T} of each side of the top-quark decay. For each scan point, one obtains m_{T}^1 and m_{T}^2 on each side of the decay and chooses the maximum of the two as $m_{\mathrm{T}}^{\mathrm{max}}$. The m_{T2} is then chosen to be the minimum of scanned $m_{\mathrm{T}}^{\mathrm{max}}$. The m_{T2} also exhibit same kinematic edge like m_{T} around the pair produced particle's mass, and for the VBF analysis, a sharp drop in m_{T2} distribution around $m_{\mathrm{T2}} = m_{\mathrm{top}}$ can be seen in Fig. 5.8. Several key BDT input variables and their modeling of data in this validation region are shown in Fig. 5.9. The WW background MC sample is shown to agree well with

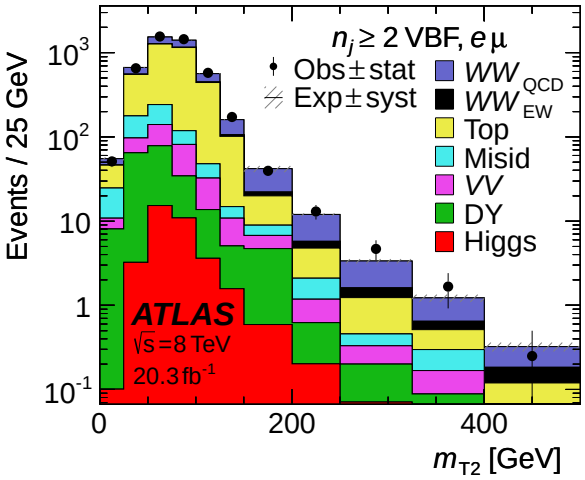


Figure 5.8: The m_{T2} distribution to define WW validation region [18]. The requirement of $m_{\mathrm{T2}} > 160$ GeV is applied to define the WW validation region.

the data.

5.4.3 ggF Higgs Boson Background

The ggF Higgs boson process is taken as a background source in this analysis. Method B is used to estimate the ggF contribution. Orthogonal selection regions used in the ggF analysis dedicated to the search of ggF produced Higgs boson are used as a control region to profile the contribution in the VBF signal region [18]. The m_{T} distribution in the ggF analysis shown in Fig. 5.10 is included in the likelihood of the global fit (but with different binning) for the VBF analysis in order to profile the normalization of ggF process. MC calculation is performed up to NLO with the POWHEG generator; kinematics of up to one emission of quark or gluon is calculated at the matrix element level. Additional emission of QCD radiation is modeled by the PYTHIA parton shower program interfaced to POWHEG.

The signal region of the analysis requires $n_j \geq 2$ and central jet veto. This restricts the events to have exclusively two jets in the central region. Such exclusive jet bin restriction on a number of jets introduces extra terms known as Sudakov double logarithms $L = \ln^2(p_{\mathrm{T}}^{cut}/Q)$ at each order of QCD perturbation expansion, where Q is the hard scale of the process and the p_{T}^{cut} is the threshold that defines the jet bins. It has been shown for ggF Higgs boson process that there is a value of p_{T}^{cut} that exactly cancels out the higher order QCD correction terms removing the dependence on Q [78]. When the exclusive jet bin cross section no longer depends on Q, the standard prescription of varying scale will not correctly estimate the unaccounted higher order calculations. The Stewart-Tackmann method [78] thus recommends estimating the uncertainty

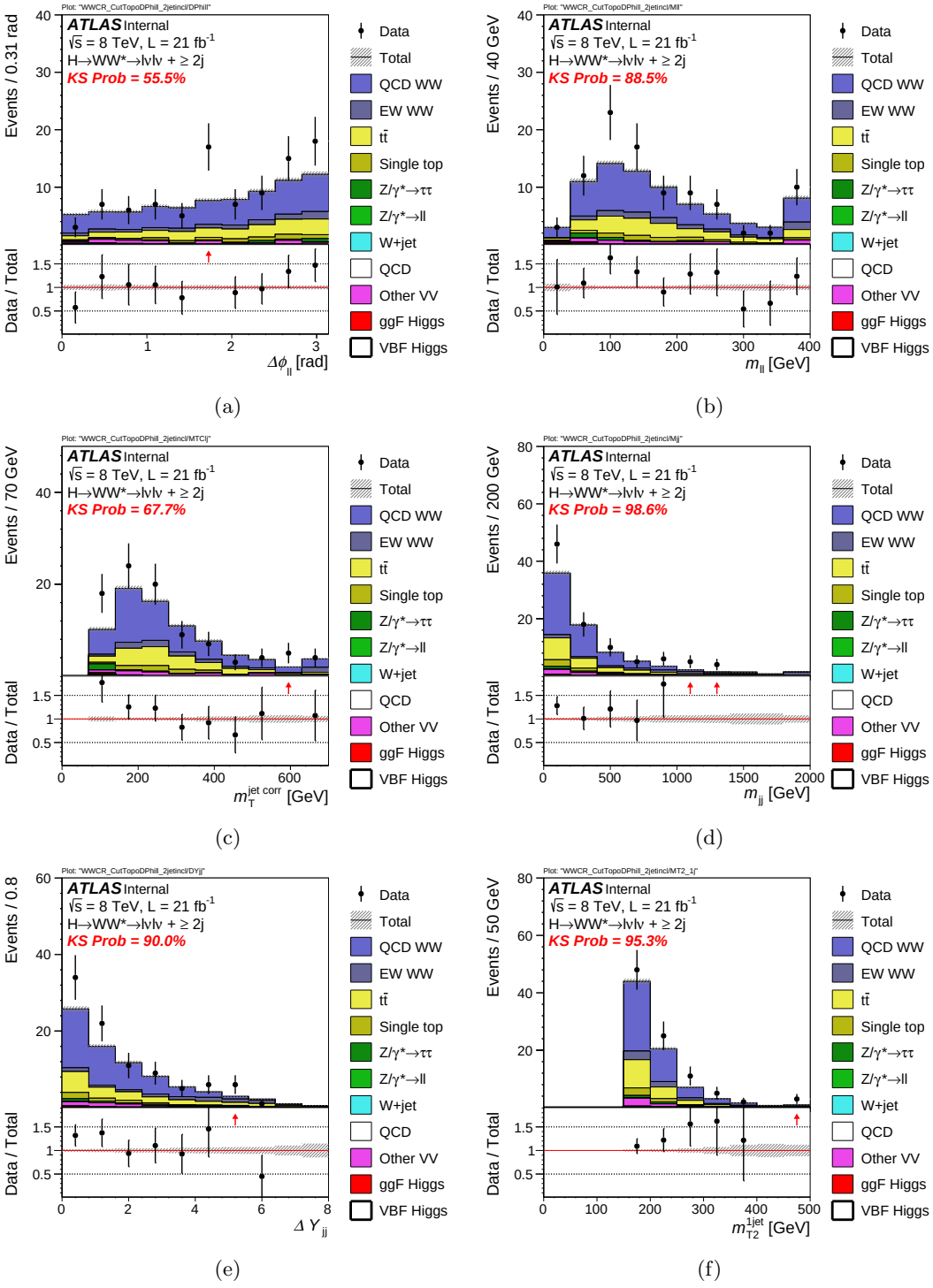


Figure 5.9: The BDT input distributions in the WW validation region. The WW background estimation relies on the MC calculation. It is crucial that the modeling of the leading background is understood. The WW validation region verifies that the MC estimate of WW agrees with data.

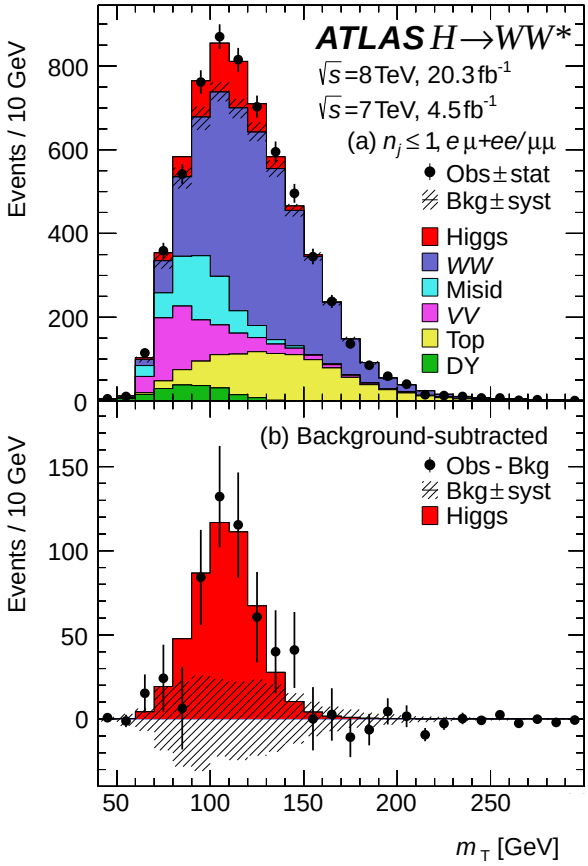


Figure 5.10: The m_{T} distribution in the ggF analysis. The binning adopted in the actual likelihood fit is different than what it shown here. Details of the binning choices can be found in Ref. [18].

Table 5.6: Summary of uncertainties in ggF Higgs boson background process.

	ΔN_{ggF} (%)		
Uncertainty source	Bin 1	Bin 2	Bin 3
Jet binning	29	29	29
Scale	7	3	48
UE/PS	15	15	15

on jet bin exclusive ggF cross section with central jet veto $\Delta(\sigma_{=2\mathrm{j}})$ as follows,

$$\Delta(\sigma_{=2\mathrm{j}}) = \Delta(\sigma_{\geq 2\mathrm{j}}) \oplus \Delta(\sigma_{\geq 3\mathrm{j}}), \tag{5.8}$$

where, $\Delta(\sigma_{\geq 2\mathrm{j}})$ is the uncertainty on the ggF cross section with $n_j \geq 2$ and $\Delta(\sigma_{\geq 3\mathrm{j}})$ is the uncertainty on the ggF cross section with explicitly requiring that there be central jets. The MCFM calculation was used to obtain the uncertainties $\Delta(\sigma_{\geq 2\mathrm{j}})$ and $\Delta(\sigma_{\geq 3\mathrm{j}})$. They are evaluated to be 27% and 13%, respectively. Another significant source of uncertainty comes from the UE/PS modeling because the second jet in the ggF process events in the signal region is modeled by PYTHIA. The difference in yields between ggF MC sample events showered by PYTHIA and HERWIG is 15% which is taken as an uncertainty. When factorization and renormalization scales are varied kinematics of the events change; the O_{BDT} distribution of a process changes and the change results in difference in acceptance. The changes in acceptance for BDT bin 1 and 2 are $3\% - 7\%$ while the difference was 48% in BDT bin 3. These are also taken as sources of uncertainty.

5.4.4 Top Quark Background

The top-quark background starts as the most dominant background in the analysis. Even after events are required to pass $n_b = 0$ selection, the prolific top-quark background remains one of the dominant backgrounds. Residual top-quark background is due to limited η coverage of the tracker, inefficiency of the b-tagging algorithm in the tracking region, and the p_{T} threshold required on the jets.

The background is estimated using Method C. The control region is defined by the same selections as the signal region except $n_b = 1$. The choice of the control region is motivated by the truth level study [79] that shows that events entering into the signal region have only one b-quark jet that has evaded the identification of b-tagging algorithm. Figs. 5.11 and 5.12 show BDT input variables in the top-quark control region. The O_{BDT} distribution in the top-quark control region is shown in Fig. 5.13. The modeling of the variables shows good agreement between data and MC prediction. No data events are observed at the BDT bin 3; for the extrapolation, BDT bin 2 and 3 are merged in the control region and merged bin extrapolates to BDT bin 2

and 3 in the signal region separately.

As the extrapolation factor α for each bin are obtained from the MC for the analysis, the appropriate theoretical uncertainties are included for the analysis. A similar assessment of the theoretical uncertainties done for the WW background is carried out for the top-quark background and shows that the generator differences yields the largest uncertainty. All other sources of uncertainties are negligible. Three generators POWHEG+ HERWIG, MC@NLO+ HERWIG, and ALPGEN+ HERWIG are compared to each other. The comparison provides the difference between NLO and multileg LO calculation as well as the differences in NLO calculations and the difference in the matching procedure to the parton shower program. The largest difference is observed between MC@NLO+ HERWIG to ALPGEN+ HERWIG. Fig. 5.14 shows the ratio of α's for the two generators with largest differences. The blue histogram is used[§] and the uncertainty of 10%, 12%, and 21% is used in BDT bin 1–3, respectively.

5.4.5 Z/γ^* Background

The $Z/\gamma^* \to \tau\tau$ background is estimated by Method B. The normalization factor is derived from a dedicated control region. In addition to the requirement listed in Table 5.1, the control region is defined by requiring $|m_{\tau\tau} - m_Z| < 25$ GeV, $m_{\ell\ell} < 80$ GeV (for the $e\mu$ sample), and $O_{\rm BDT} > -0.48$. The additional criteria increase purity on $Z/\gamma^* \to \tau\tau$ process and the $O_{\rm BDT}$ requirement ensures that the events in the control region are kinematically similar to the ones in the signal region. The normalization factor $\bar{N}_{\rm CR}/\bar{B}_{\rm CR}$ as defined in the Sec. 5.4.1 is 0.9 ± 0.3 The statistical uncertainty from the control region is included in the final fit. The Fig. 5.15 shows the distribution of the $m_{\tau\tau}$ in the control region.

The $Z/\gamma^* \to ee$, $\mu\mu$ background in the $ee/\mu\mu$ is the leading background source. The background is estimated using Method D. A control region with low missing transverse momentum is used to model the $Z/\gamma^* \to ee$, $\mu\mu$ contribution in the $ee/\mu\mu$ sample. The CR for the $Z/\gamma^* \to ee$, $\mu\mu$ background estimation is defined to have all the same selections as the signal region except the missing transverse momentum requirements are replaced by requiring 25 GeV $< E_{\rm T}^{\rm miss} <$ 45 GeV. The CR$'$ (SR$'$) is defined to be the same as the CR (SR) except the requirement $m_{\ell\ell} < m_Z - 15$ GeV is replaced by $|m_{\ell\ell} - m_Z| < 15$ GeV. The Fig. 5.16 shows the distribution of $O_{\rm BDT}$ in the CR. The binning shown in the figure is the same used in the signal region fit. The BDT bin 2 and 3 are merged during the extrapolation in order to reduce statistical uncertainty. A correction factor of $f = 0.83 \pm 0.22$ is obtained from the MC sample.

[§]The red histogram used the fully reconstructed objects in determining $O_{\rm BDT}$ and due to the low statistics generator level comparison was used.

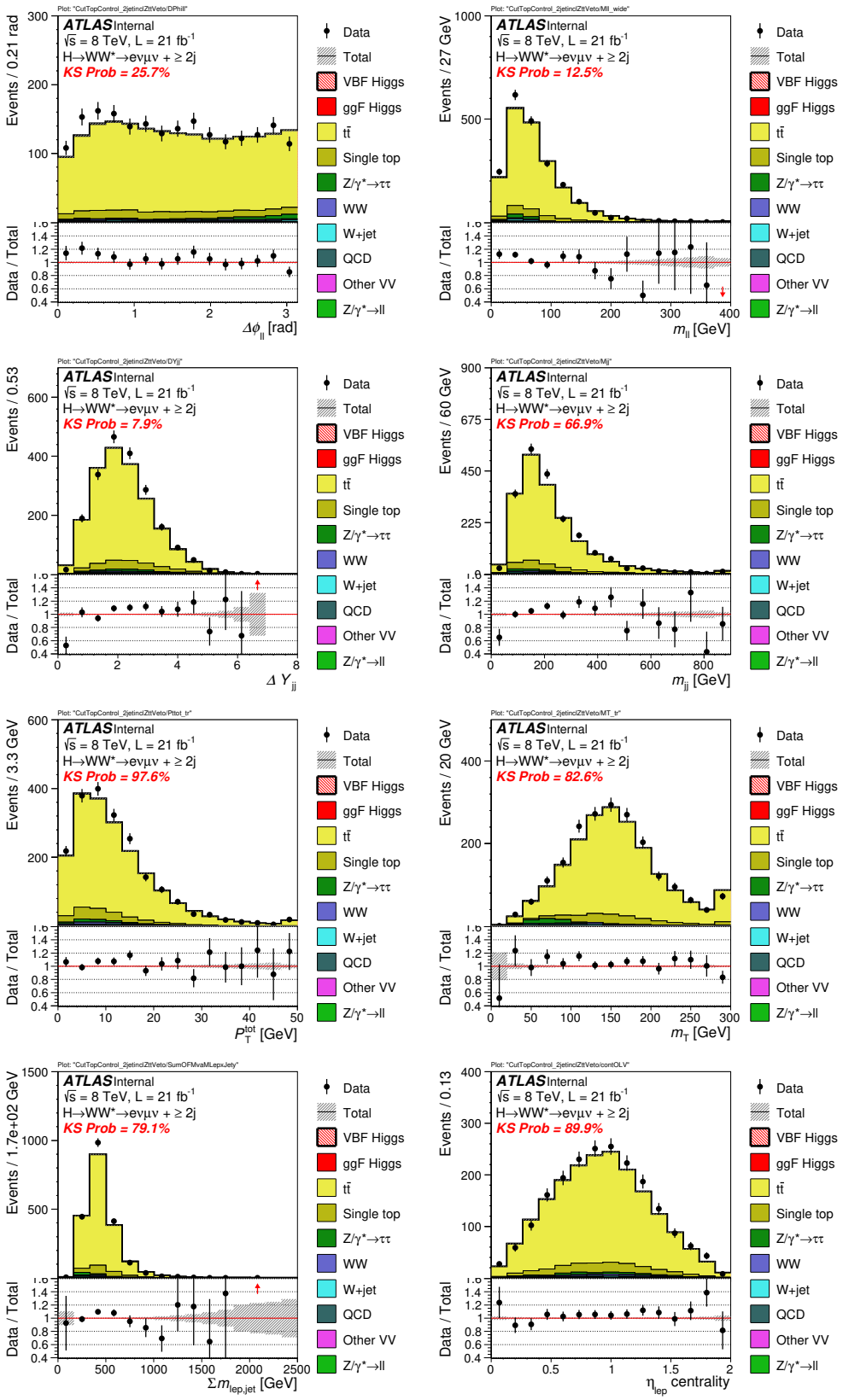


Figure 5.11: Distributions of $\Delta\phi_{\ell\ell}$, $m_{\ell\ell}$, Δy_{jj}, m_{jj}, $p_{\mathrm{T}}^{\mathrm{sum}}$, m_{T}, $\Sigma\, m_{\ell j}$, and $\sum C_{\ell}$ in the $e\mu$ top-quark background control region.

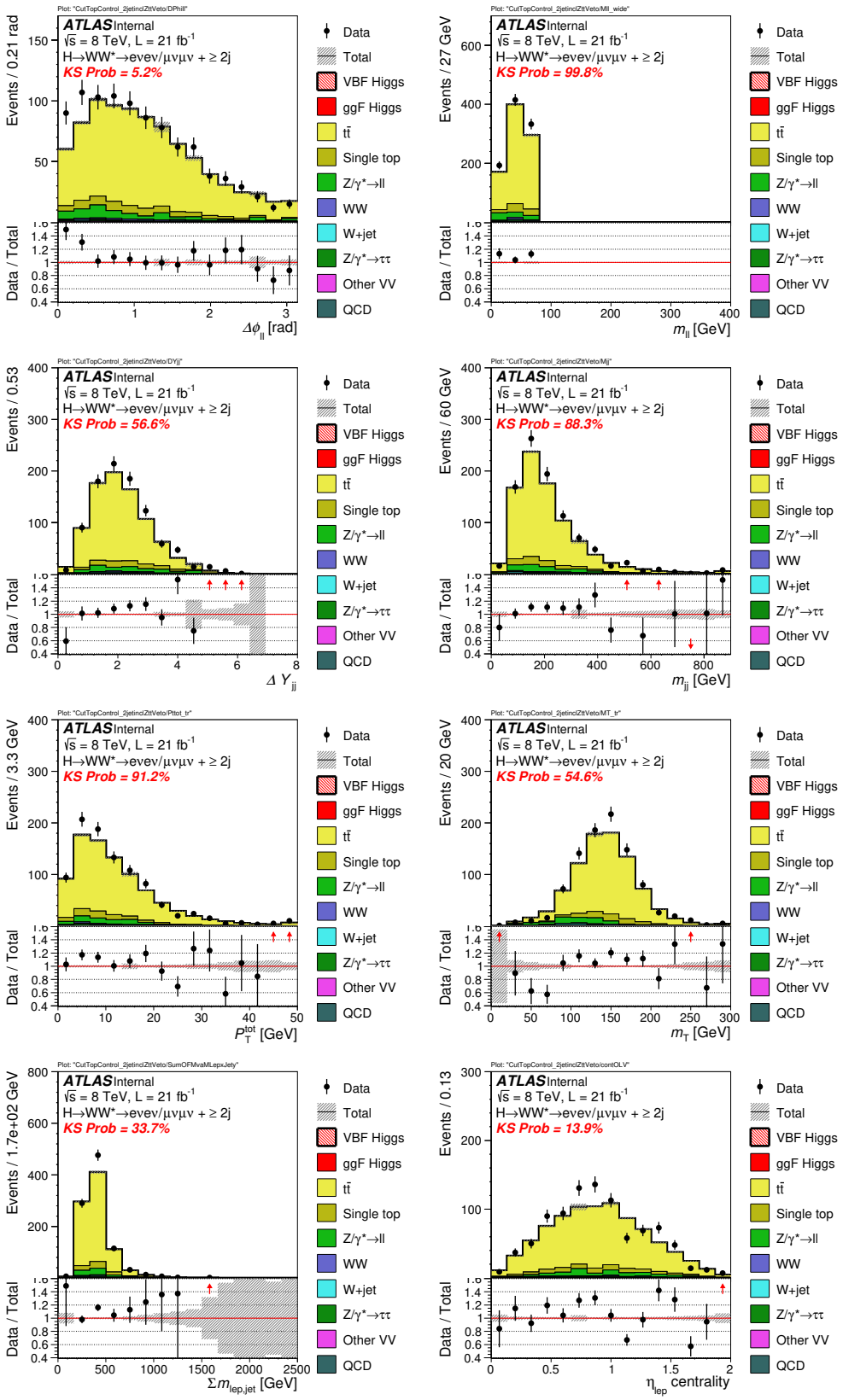


Figure 5.12: Distributions of $\Delta\phi_{\ell\ell}$, $m_{\ell\ell}$, Δy_{jj}, m_{jj}, $p_{\mathrm{T}}^{\mathrm{sum}}$, m_{T}, $\Sigma\, m_{\ell j}$, and $\sum C_{\ell}$ in the $ee/\mu\mu$ top-quark background control region.

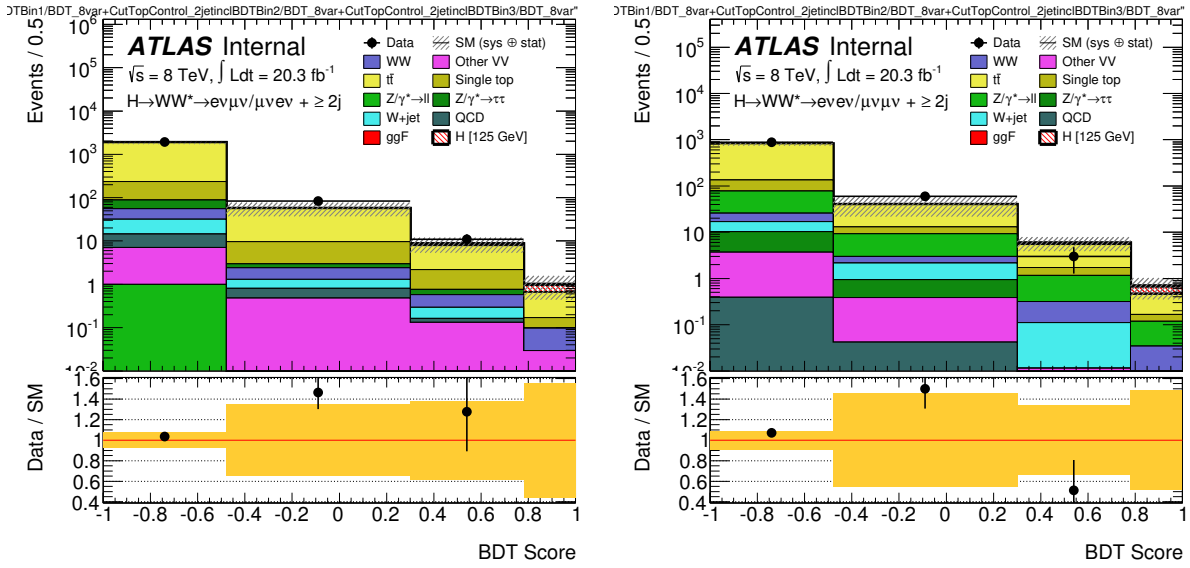


Figure 5.13: Distributions of O_{BDT} score in the top-quark control region after all preselection, for $e\mu$ (left) and $ee/\mu\mu$ (right) samples. The bin boundaries are the same as in the fit. In order to illustrate the consistency between the trend observed in the NFs and the understanding of the theoretical modeling, coverage of the data/MC ratio by the estimated theoretical uncertainties on the top yield in the CR was calculated and included in the yellow band. The values are 5%, 35%, 38% and 75% respectively for all four BDT bins. Note that these uncertainties are not used in the fit, there they are replaced by the extrapolation uncertainties.

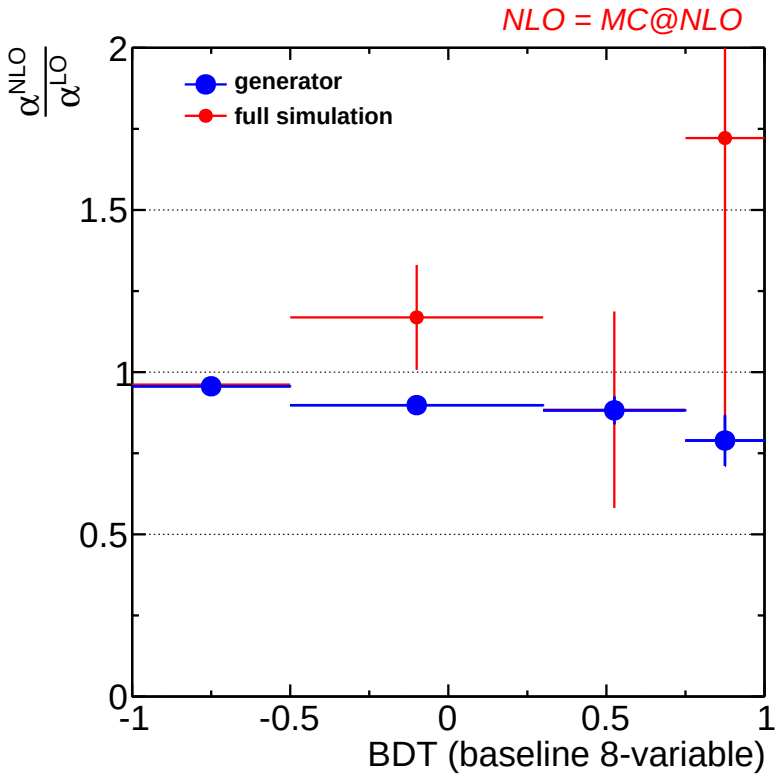


Figure 5.14: Difference in the α for each BDT bins using POWHEG+ HERWIG to ALPGEN+ HERWIG [80].

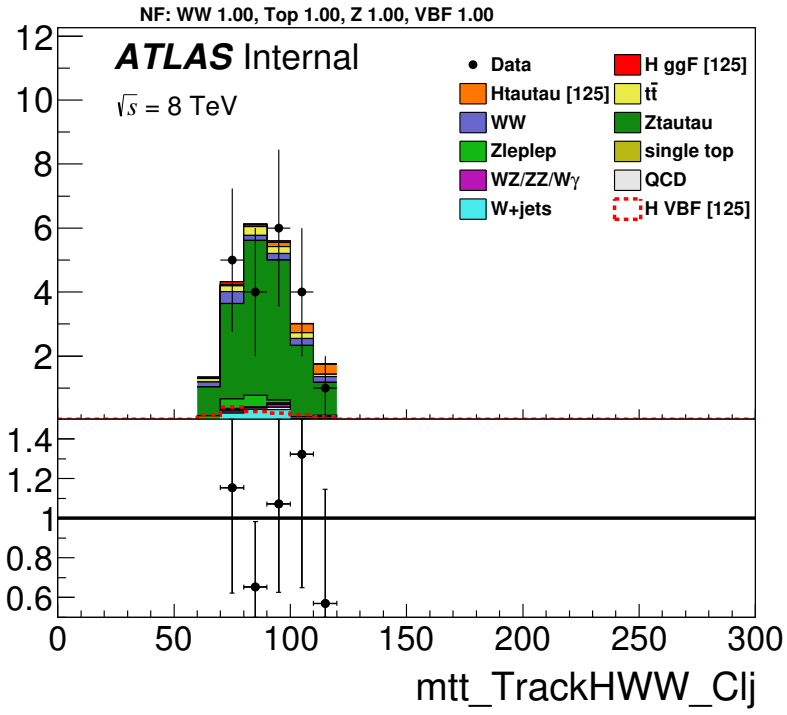


Figure 5.15: The $m_{\tau\tau}$ distribution in the $Z/\gamma^* \to \tau\tau$ control region. In addition to the requirement listed in Table 5.1, the control region is defined by requiring $|m_{\tau\tau} - m_Z| < 25$ GeV, $m_{\ell\ell} < 80$ GeV (for the $e\mu$ sample), and $O_{\text{BDT}} > -0.48$. The normalization factor $\bar{N}_{\text{CR}}/\bar{B}_{\text{CR}} = 0.9 \pm 0.3$ is obtained.

5.5 Systematic Uncertainties

A summary of the leading systematic uncertainties is presented in Table 5.7. The relative size of the uncertainties are quoted by their impact to the signal strength measurement.

In additions to the theoretical source of uncertainties which have been mostly discussed in the previous section Sec. 5.4, there are systematic uncertainties arising from the experiment. The experimental uncertainties are mainly related to the object reconstructions. Various parameters used during the object reconstructions are systematically varied and the propagated changes in the O_{BDT} distribution are taken as uncertainties and correlated across the BDT bins.

The uncertainties from lepton reconstructions are evaluated to have less than $2-3\%$ effect. Uncertainties mainly arise from the uncertainty on the particle identification efficiencies. They are around 1% for both the electron and muon reconstruction and are correlated across BDT bins. The uncertainty from calibrations of around $\leq 0.5\%$ is also considered. The impact of the uncertainties from the lepton reconstruction, summarized in the row "Muons and electrons" in Table 5.7, is found to be $2-3\%$.

The sizes of uncertainties related to the jet energy scale (JES) and jet energy resolution (JER) calibration are much larger. The basics of the jet energy scale calibration comes from using a tag-and-probe method

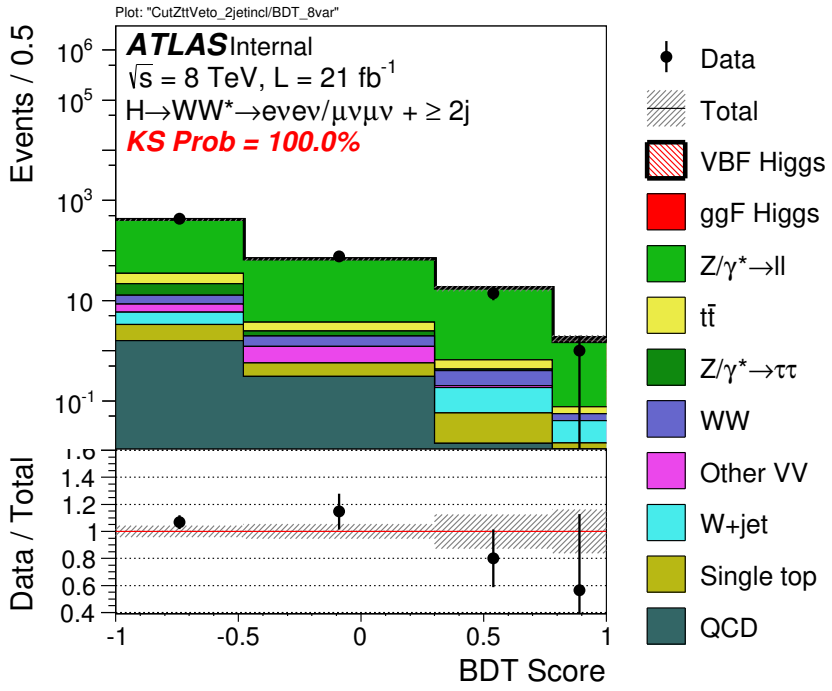


Figure 5.16: Distribution of O_{BDT} in the low missing transverse momentum region. The size of $Z/\gamma^* \to ee$, $\mu\mu$ background obtained from this distribution is extrapolated using the extrapolation factor α obtained from the region with $|m_{\ell\ell} - m_Z| < 15$ GeV. A correction obtained from the MC calculation on the difference between events from $m_{\ell\ell} < m_Z - 15$ GeV and $|m_{\ell\ell} - m_Z| < 15$ GeV is applied to the extrapolation.

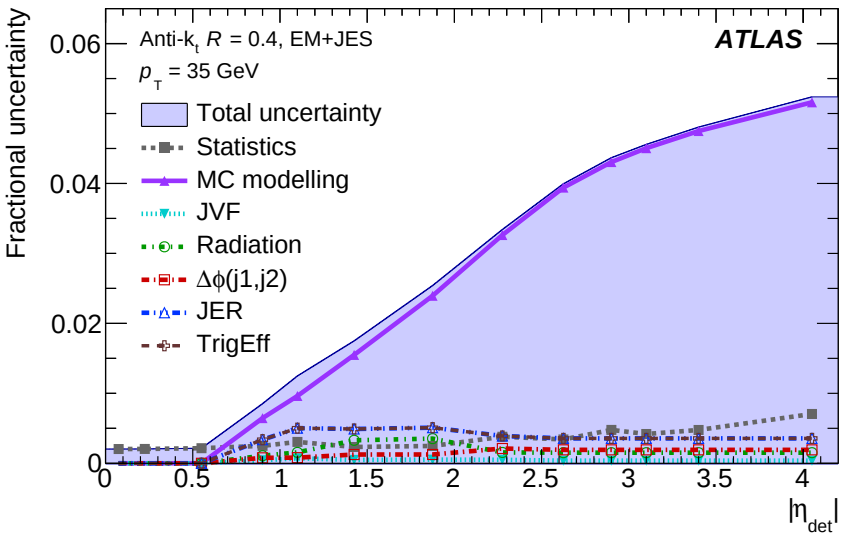


Figure 5.17: Breakdown of jet energy scale calibration uncertainty. The leading contribution of the uncertainties for jets with high $|\eta_{\text{det}}|$ is the one labeled "MC modeling". The MC modeling uncertainty is from the prediction difference between PYTHIA and HERWIG generators. The VBF analysis signal region prefers jets with high $|\eta|$; the MC modeling uncertainty from the JES for high $|\eta|$ impacts the analysis significantly [45].

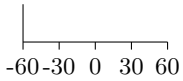
from dijet events. Due to inefficiencies in data-taking for dijet events with high $|\eta|$ jets, MC sample is used to extrapolate calibration scales from data events with low $|\eta|$ jets to events with $|\eta|$ jets [45]. The extrapolated calibration scales are used to calibrate jets with high $|\eta|$. The difference between PYTHIA to HERWIG in the prediction of the extrapolation from low $|\eta|$ region to high $|\eta|$ region results in significant size in the uncertainty. The Fig. 5.17 shows the uncertainty breakdown for the JES calibration. The dominant source of the uncertainty comes from the "MC modeling", which is corresponding to the difference between PYTHIA and HERWIG predictions. The impact of the JES and JER uncertainties to the VBF analysis is summarized in the row "Jets" in Table 5.7, which is $11-15\%$.

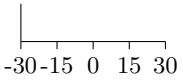
The leading source of theoretical uncertainty comes from the modeling of the VBF process. The calculation of VBF process comes with a relatively large theoretical uncertainty despite the nature of the process is purely an electroweak interaction (cf. Ref. [108]). The large uncertainty arise from the parton shower modeling of VBF jets and the underlying event modeling. A comparison between MC samples of VBF produced Higgs boson process showered by PYTHIA and HERWIG results in acceptance difference of [-0.2%, 1.8%, 11%] for the BDT bin 1–3, respectively. The increase in the uncertainty can be understood as coming from $p_{\text{T}}^{\text{sum}}$ variable, which are sensitive to the soft jet activities.

The impact of theoretical and experimental systematic uncertainty on the signal strength measurement is found to be $15-22\%$ and $14-18\%$, respectively. In addition, the uncertainty on the measured integrated luminosity impacts the signal strength measurement by $3-5\%$.

Table 5.7: The table and caption are reproduced from Ref. [18]. Summary of uncertainties on the signal strength μ. The table gives the relative uncertainties for VBF production (left), and ggF production (right). The "profiled signal region" indicates the contribution of the uncertainty on the ggF signal yield to the μ_{VBF} measurement and vice versa.

Source	Observed $\mu_{\text{VBF}} = 1.27$			Observed $\mu_{\text{ggF}} = 1.02$		
	Error +	Error −	Plot of error (scaled by 100)	Error +	Error −	Plot of error (scaled by 100)
Data statistics	0.44	0.40		0.19	0.19	
Signal regions	0.38	0.35		0.14	0.14	
Profiled control regions	0.21	0.18		0.12	0.12	
Profiled signal regions	0.09	0.08		0.03	0.03	
MC statistics	0.05	0.05		0.06	0.06	
Theoretical systematics	0.22	0.15		0.19	0.16	
Signal $H \to WW^*$ $\mathcal{B}$	0.07	0.04		0.05	0.03	
Signal ggF cross section	0.03	0.03		0.13	0.09	
Signal ggF acceptance	0.07	0.07		0.06	0.05	
Signal VBF cross section	0.07	0.04		-	-	-
Signal VBF acceptance	0.15	0.08		-	-	-
Background WW	0.07	0.07		0.08	0.08	
Background top-quark	0.06	0.06		0.04	0.04	
Background misid. factor	0.02	0.02		0.06	0.06	
Others	0.03	0.03		0.02	0.02	
Experimental systematics	0.18	0.14		0.08	0.08	
Background misid. factor	0.02	0.01		0.04	0.04	
Bkg. $Z/\gamma^* \to ee$, $\mu\mu$	0.01	0.01		0.03	0.03	
Muons and electrons	0.03	0.02		0.05	0.04	
Missing transv. momentum	0.05	0.05		0.02	0.01	
Jets	0.15	0.11		0.03	0.03	
Others	0.06	0.06		0.03	0.03	
Integrated luminosity	0.05	0.03		0.03	0.02	
Total	0.53	0.45		0.29	0.26	





5.6 Fit Procedures

The O_{BDT} distributions are fitted in the signal region and control regions simultaneously to extract best-fit signal strength μ_{VBF}. The likelihood function definition for the global fit is defined as follows,

$$\begin{aligned}\mathcal{L} = & \underbrace{\prod_i f\left(N_i^{\mathrm{SR}}|\mu\cdot S_i^{\mathrm{SR}}\prod_k \nu_{ik}(\theta_k) + \sum_b \beta_b \cdot B_{bi}^{\mathrm{SR}}\cdot\prod_l \nu_{bl}(\theta_l)\right)}_{\text{Poisson for SR with signal strength } \mu \text{ and prediction } S \text{ and } B} \\ & \times \underbrace{\prod_{j,m} f\left(N_j^{\mathrm{CR}_m}|\sum_c \beta_c \cdot B_{cj}^{\mathrm{CR}_m}\cdot\prod_s \nu_{cs}(\theta_s)\right)}_{\text{Poisson for CR to profile background source of } c} \\ & \times \underbrace{\prod_t g\left(\vartheta_t|\theta_t\right)}_{\text{Gauss. for syst.}}\,, \end{aligned} \tag{5.9}$$

where $\{i,j\}$ runs over the BDT bins, m runs over the control regions, $\{t,l,k,s\}$ runs over the systematics, and $\{b,c\}$ runs over background sources.

The function $f(N|\lambda) \equiv e^{-\lambda}\lambda^N/N!$ is the Poisson function of the probability of observing N events given λ expected events. The expected value of λ in the first term is of the form $\mu\cdot S + \sum_k B_k$ where S and B_i are the signal and background contributions. The contributions S and B_i in the Eq. (5.9) have response function ν multiplied to parameterize the effect of systematic uncertainties on the signal and background contributions. The θ's are the parameters representing the systematic sources.

The second Poisson function represents the profiling of the background contributions in different control regions. The profiled β values from the second Poisson term will constrain the background contributions in the signal region (i.e. the first Poisson term). The same systematic effects are also multiplied in the expected value λ of the Poisson function.

The last term constrains the parameters representing the systematic sources. Each term has a Gaussian function of the form $g(\vartheta|\theta) \equiv e^{-(\vartheta-\theta)^2/2}/\sqrt{2\pi}$, where the ϑ is the central value of the systematic uncertainty source and the θ is the nuisance parameters associated for the systematic uncertainty sources.

To test the null hypothesis that there is no VBF process in the events (i.e. $R \equiv \mu_{\mathrm{VBF}}/\mu_{\mathrm{ggF}} = 0$) a test statistics of the following is used,

$$q_R = -2\ln\left.\frac{\mathcal{L}(R,\vec{\theta})}{\mathcal{L}_{\max}}\right|_{\vec{\theta}=\hat{\vec{\theta}}_R}\,, \tag{5.10}$$

where the $\mathcal{L}_{\max}$ is the maximum value over all possible $\vec{\theta}$ and R values, and $\hat{\vec{\theta}}_R$ is the value of $\vec{\theta}$ that maximize the $\mathcal{L}$ for a given R. To evaluate p-value of the null hypothesis consider the probability density

function pdf($q_R|R=0$), where $R=0$ represents the assumption of the null hypothesis of expecting no VBF contributions in data. The p-value is then defined to be,

$$p_0 = \int_{q_0^{obs}}^{\infty} \mathrm{pdf}(q_0|0)dq_0. \tag{5.11}$$

The signficance is then taken as the one-sided tail of a Gaussian distribution, $Z_0 = \sqrt{2}\mathrm{erf}^{-1}(1-2p_0)$.

5.7 Results

As the VBF analysis searching for the VBF produced Higgs boson in its WW^* decay, the mass of the Higgs boson is set to be $m_H = 125.36$ GeV from the measurement from $H \to ZZ^*$ and $H \to \gamma\gamma$ decay channels [81]. By fixing he m_H value, the size and shape of the signal processes are determined. The Fig. 5.18 shows the post-fit distribution of the O_{BDT} and m_{T} distributions after the obtained nuisance parameter values from the global fit described in Sec. 5.6 are applied appropriately. A clear presence of VBF process can be seen in the $e\mu$ sample shown in Fig. 5.18(a), while the $ee/\mu\mu$ sample shows moderate excess. The m_{T} distribution from selection-based cross-check analysis is shown in Fig. 5.19(a). A result consistent with the signal hypothesis can also be seen. The Fig. 5.19(b) shows a scattered plot of observed data in the variables m_{T} and m_{jj}. The scatter plot shows that observed events populate phase-spaces where the signals are expected to be. The Table 5.8 shows the summary of each contribution broken down by BDT bins for both 7 and 8 TeVanalysis. Uncertainty on each yield includes all uncertainties from every source.

The scan of test statistics q_R is shown in Fig. 5.20. The y-axis refers to the test statistics where Λ represents the argument of the logarithm in Eq. (5.10). The resulting best-fit value for the ratio is

$$\frac{\mu_{\mathrm{VBF}}}{\mu_{\mathrm{ggF}}} = 1.26^{+0.61}_{-0.45}(\mathrm{stat})^{+0.50}_{-0.26}(\mathrm{syst}) = 1.26^{+0.79}_{-0.53}. \tag{5.12}$$

The $-2\ln\Lambda$ value at $\mu_{\mathrm{VBF}}/\mu_{\mathrm{ggF}} = 0$ is interpreted as the significance of the existence of VBF process. The observed significance is 3.2 standard deviations; the expected significance is 2.7 standard deviations. This establishes the first evidence for the VBF production mode in the $H \to WW^* \to \ell\nu\ell\nu$ final state. The same significance for the VBF production in selection-based cross-check analysis is obtained and the observed and expected significance is 3.0 and 2.1 standard deviations, respectively. The compatibility of the two analysis is checked in 8 TeV analysis by generating pseudoexperiments with only the statistical uncertainties considered for simplicity and found to be in good agreement [18].

Table 5.8: Signal region yields with uncertainties reproduced from Ref. [18]. The tables give the post-fit yields separated for the 8 and 7 TeV data analyses. The N_{signal} columns show the expected signal yields from the ggF and VBF production modes, with values scaled to the observed combined signal strength.The yields and the uncertainties take into account the pulls and data-constraints of the nuisance parameters, and the correlations between the channels and the background categories. The quoted uncertainties include the theoretical and experimental systematic sources and those due to sample statistics. Values less than 0.1 (0.01) events are written as 0.0 (-).

	Summary				Composition of N_{bkg}				
Channel	N_{obs}	N_{bkg}	N_{VBF}	N_{ggF}	N_{WW}	N_{top}	N_{misid}	N_{VV}	N_{DY}
(a) 8 TeV data									
$e\mu$ sample									
bin 1	37	36 $\pm$4	4.9 ± 0.5	3.3 ± 1.2	5.0 ± 1.5	19 $\pm$2.7	5.5 ± 2.5	2.3 ± 0.7	3.6 ± 1.5
bin 2	14	6.5 ± 1.3	4.9 ± 0.5	1.4 ± 0.5	1.7 ± 0.7	2.3 ± 1.1	0.7 ± 0.3	0.7 ± 0.2	0.6 ± 0.2
bin 3	6	1.2 ± 0.3	3.8 ± 0.7	0.4 ± 0.3	0.3 ± 0.1	0.4 ± 0.1	-	0.1 ± 0.0	0.2 ± 0.1
$ee/\mu\mu$ sample									
bin 1	53	46 $\pm$6	2.6 ± 0.3	1.7 ± 0.6	3.1 ± 1.0	12 $\pm$1.6	1.1 ± 0.2	1.0 ± 0.3	28 $\pm$5
bin 2	14	8.4 ± 1.8	3.0 ± 0.4	0.7 ± 0.3	0.9 ± 0.3	1.5 ± 0.5	0.2 ± 0.1	0.3 ± 0.1	5.2 ± 1.7
bin 3	6	1.1 ± 0.4	2.1 ± 0.4	0.2 ± 0.2	0.1 ± 0.1	0.3 ± 0.1	-	-	0.5 ± 0.3
(a) 7 TeV data									
$e\mu$ sample									
bin 1	6	3.0 ± 0.9	0.6 ± 0.1	0.4 ± 0.2	0.7 ± 0.2	1.0 ± 0.5	0.1 ± 0.0	0.3 ± 0.1	0.8 ± 0.6
bin 2–3	0	0.7 ± 0.2	1.1 ± 0.1	0.2 ± 0.1	0.2 ± 0.1	0.3 ± 0.2	-	-	-
$ee/\mu\mu$ sample									
bin 1–3	3	4.1 ± 1.3	1.0 ± 0.1	0.3 ± 0.1	0.6 ± 0.2	0.7 ± 0.3	-	0.2 ± 0.1	2.5 ± 1.1

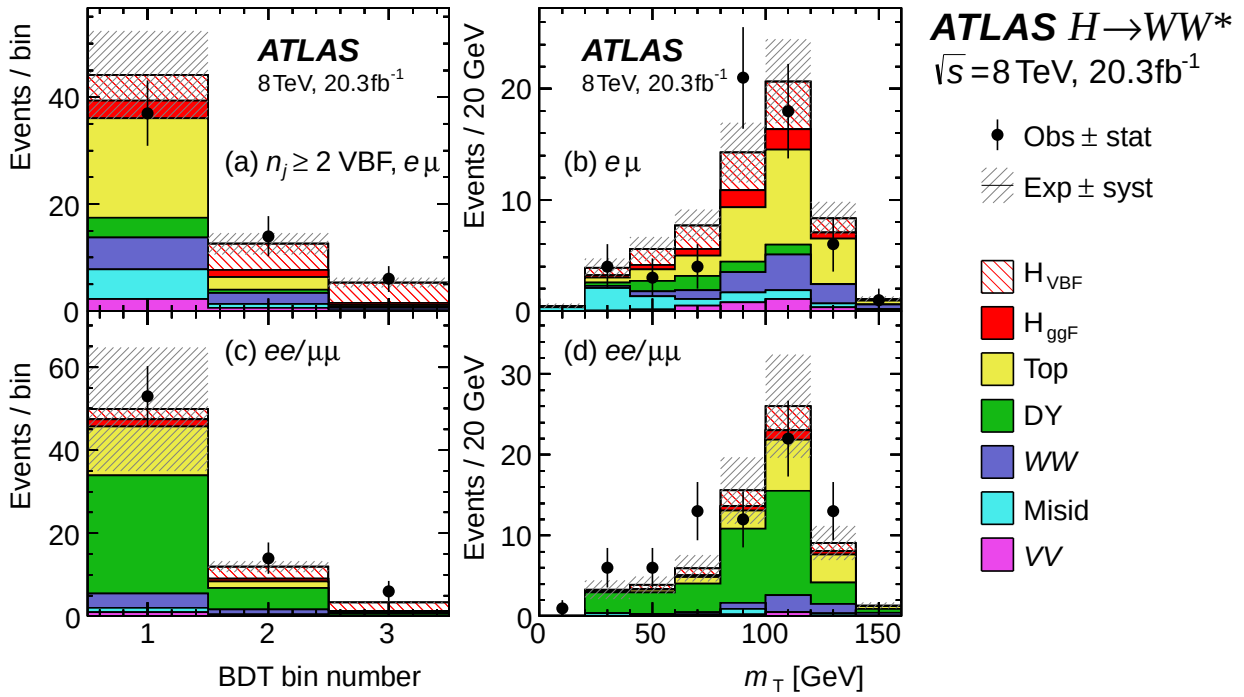


Figure 5.18: Post-fit O_{BDT} and m_{T} distribution for the BDT analysis [18]. (a) O_{BDT} and (b) m_{T} distribution in $e\mu$ sample is shown. (c) O_{BDT} and (d) m_{T} distribution in $ee/\mu\mu$ sample is shown.

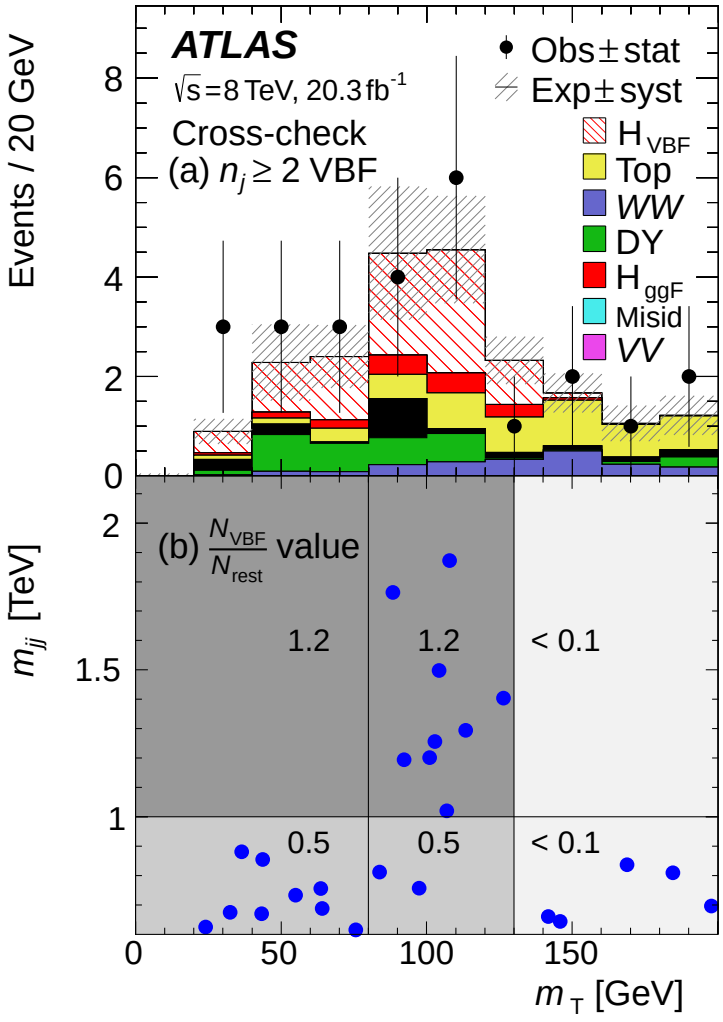


Figure 5.19: (a) Post-fit m_{T} distribution for the cross-check analysis [18]. (b) Scatter plot of m_{jj} vs. m_{T} for data. For each bin, the ratio $N_{\mathrm{VBF}}/N_{\mathrm{rest}}$ is stated in the plot, where N_{rest} includes all processes other than the VBF signal.

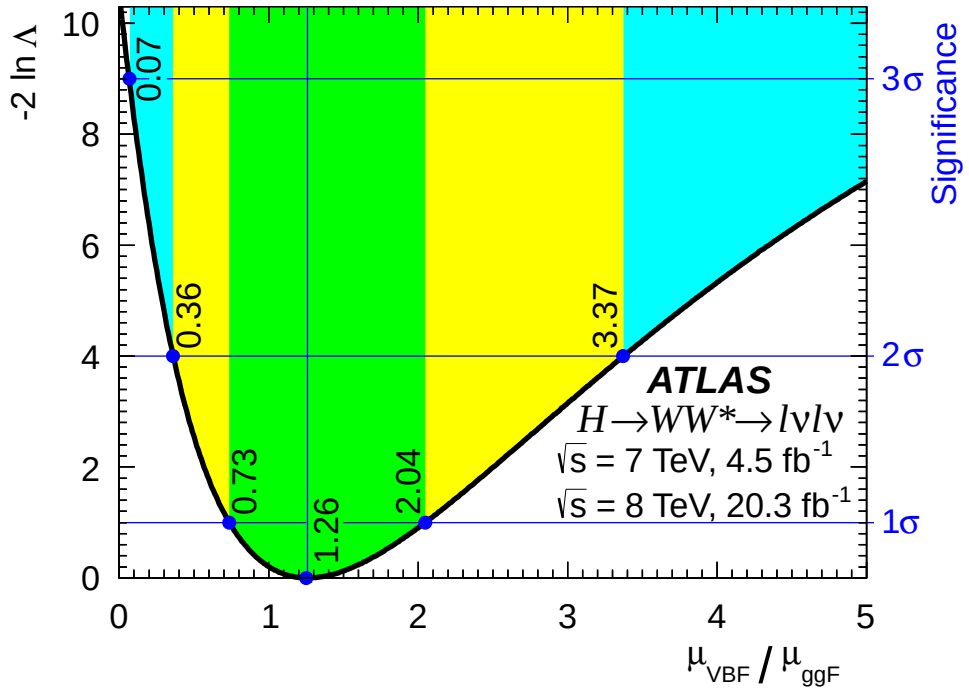


Figure 5.20: Likelihood scan as a function of $\mu_{\mathrm{VBF}}/\mu_{\mathrm{ggF}}$ for $m_H = 125.36$ GeV [18]. The value of the likelihood at $\mu_{\mathrm{VBF}}/\mu_{\mathrm{ggF}} = 0$ gives the significance of the VBF signal at 3.2 standard deviations. The inner (middle) [outer] band shaded darker (lighter) [darker] represents the one (two) [three] standard deviation uncertainty around the central value represented by the vertical line.

Chapter 6

Search for Higgs Boson Decay to Additional Scalar Bosons

This chapter describes a search for the Higgs boson decaying to (pseudo)scalar pairs in the $bb\mu\mu$ final state at $\sqrt{s} = 13$ TeV with the ATLAS detector at the LHC. The discovery of the Higgs boson has opened up a new avenue to search for physics beyond the SM. One way to use the Higgs boson as a new tool for discovery is to search for non-SM decays of the Higgs boson. Such searches provide a unique window onto light hidden sectors, consisting of particles neutral under SM gauge groups as explained in Sec. 2.5.

6.1 Theoretical Motivation

In the SM, a Higgs boson with a mass of 125 GeV has a tiny width. This can be understood by taking a closer look at the SM Higgs branching ratios, which are shown in Fig. 6.1. In Fig. 6.1, the colored contributions are the SM decays and the gradient black-and-white contribution shows the current upper limit on the $\mathcal{B}(H \to BSM) = 34\%$ [22]. The dominant decay is to the $b\bar{b}$ final state, followed by $H \to VV$. What is common among all SM Higgs decay branching ratios is that each decay is suppressed for various reasons and this results in a tiny total Higgs width of 4 MeV [82]. The small width implies that if the Higgs boson has a non-SM coupling, even a tiny one, it can easily alter the branching ratio of the decay modes of the Higgs boson by $\mathcal{O}(10\%)$ [30]. Since the projected 95% confidence limit on $\mathcal{B}(H \to BSM)$ with 3 ab^{-1} is at the order of $5 - 10\%$ [83], without directly searching for non-SM decay modes, many models with small couplings to the Higgs boson will be left unconstrained throughout the LHC program. It is also important to note that if a non-SM decay mode exists with $\mathcal{B}$ of $5 - 10\%$, then millions of such events will be produced in 13 TeV pp collisions at the LHC, viz. 1.5 M with $\mathcal{B} = 10\%$ and 300 fb^{-1} providing ample statistics for discovery and making it an urgent question to be answered during the 13 TeV LHC running.

Of the possible non-SM decays of the Higgs boson, the possibility of Higgs boson decaying to pairs of (pseudo)scalar boson $H \to ss$ is particularly interesting. As explained in Sec. 2.5, the possibility of a Higgs boson decaying to other light scalar bosons is theoretically well motivated. In particular, models with pseudoscalars have been proposed as possible explanations for Dark Matter (DM) annhiliations producing

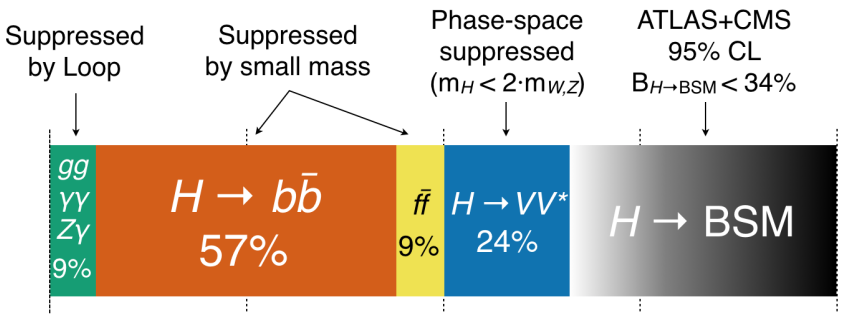


Figure 6.1: Branching ratios for the Higgs boson with $m_H = 125$ GeV. The largest is to the $b\bar{b}$ with 57%. The reasons for why the coupling strengths are small for each decays are annotated.

gamma ray excess observed by the Fermi-LAT [84]. If confirmed, it will open up an exciting possibility of studying the nature of DM using the Higgs boson as a tool. This thesis will focus on the benchmark model with pseudoscalar for the DM motivation. The search does not utilize the angular distributions to separate signal from background process. The phenomenology is the same for the cases with pseudoscalar or scalar, except angular distributions, the search will be sensitive to both cases equally. Please note that throughout the chapter the hypothesized (pseudo)scalar being searched will be represented as s and will be referred to as simply the "scalar", unless noted otherwise.

Once produced, the a can decay back to other light SM particles. As the a decays to the SM particles through its coupling to the Higgs doublet of the SM, the branching ratios of a decay to SM particles inherits from the Higgs boson decay to SM particles [85]. This means that generally the dominant decay is to the heaviest SM fermion state available by the kinematic phase-space. The branching ratios for a benchmark model of type-III Two Higgs Double Model with additional singlet pseudoscalar (2HDM+S) is shown in Fig. 6.2 [85]. Fig. 6.2 shows the hierarchical branching ratio structure with larger branching ratios to heavier SM fermion states. The largest branching ratio is thus to the $H \to ss \to 4b$ decay or the $H \to ss \to 4\tau$ decay, but such final states are contaminated by large QCD background contributions. The $H \to ss \to 4\mu$ provides a clean channel, but the branching ratio is tiny (i.e. 4×10^{-6}). The $H \to ss \to bb\mu\mu$ is a good compromise between a clean channel and appreciable branching ratio.

6.2 Data Samples

The final state of interest in this thesis is two muons plus two b-jets. This analysis uses events recorded by triggers requiring a single muon. The single muon trigger during 2015 and 2016 requiring $p_{\mathrm{T}} \geq 26$ GeV is used. Recorded data events are subjected to quality criteria, which require relevant detector components were

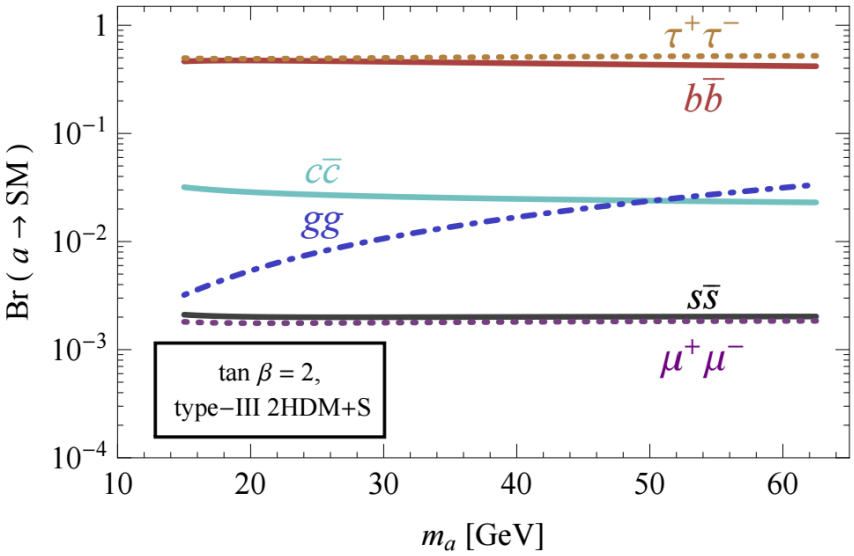


Figure 6.2: Branching ratios for the pseudoscalar a in a type-III 2HDM+S model with $\tan\beta = 2$ as a function of the pseudoscalar mass m_a [85].

operating correctly at the time of the event. After the data quality requirement, the integrated luminosity of pp collision events is 36 fb^{-1} at $\sqrt{s} = 13.5$ TeV for the 2015 and 2016 data taking combined.

6.2.1 Signal Benchmark Model

The signal samples are simulated using POWHEG to first generate $pp \to H$ process, and the events are interfaced to PYTHIA8 to simulate the $H \to ss \to bb\mu\mu$ decay. Five mass points are simulated ranging from $m_s = 20 - 60$ GeV in steps of 10 GeV. Additional pp collisions generated with PYTHIA8 are overlaid to model the effects of the pileup for all simulated events. All simulated events are processed with the same reconstruction algorithm as used for data. Samples are processed through the full ATLAS detector simulation [91] based on GEANT4 [53].

The signal process histograms are normalized to a benchmark model from Ref. [85]. The benchmark model is the type-III 2HDM+S with $\tan\beta = 2$, where the branching ratios are shown in Fig. 6.2. In this model, the branching ratio of the s pair to $bb\mu\mu$ is $\mathcal{B}(ss \to bb\mu\mu) \simeq 1.6 \times 10^{-4}$. The total cross section of the signal process assuming $\mathcal{B}(H \to ss) = 100\%$ is then,

$$\sigma_H \cdot \mathcal{B}(H \to ss) \cdot \mathcal{B}(ss \to bb\mu\mu) = 0.08 \text{ pb}, \tag{6.1}$$

where the σ_H includes the recent next-to-next-to-next-to-leading order QCD calculations with NLO electroweak corrections on ggF process [82, 86–90].

6.2.2 Preselection

A collection of loose selection criteria are required on the events. They will be referred to as preselection throughout this chapter. The preselection criteria are as follows. Events with exactly two muons with $p_{\mathrm{T}} > 7$ GeV are selected. The leading muon (μ_1) is further required to have $p_{\mathrm{T}} \geq 27$ GeV. As the s is a neutral particle, the two reconstructed muons are required to have opposite charge. Events are required to have an invariant mass $m_{\mu\mu} > 12$ GeV in order to remove low-mass meson resonances. To remove events with $Z \to \mu\mu$ process, $m_{\mu\mu} < 80$ GeV is applied. The final state requires at least two jets, so $n_j \geq 2$ is required. The preselection criteria are summarized in Table. 6.1.

6.2.3 bb and jj Samples

Events that pass the preselection criteria are split into two samples: bb sample and jj sample. The two samples are split by the criteria $n_b \geq 2$. The bb sample passes the $n_b \geq 2$, while jj sample does not. The two leading b-jets in the bb sample are called the "signal jets". The two leading jets in the jj sample are called the "proxy jets". The purpose of the proxy jets is to mimic the signal jets in the bb sample. As the b-tagging algorithm requires the use of tracking information, the signal jets in the bb sample require $|\eta_b| < 2.5$ due to limited η coverage of the ID. For consistency with the signal jets requirement, the proxy jets are required to be $|\eta_j| < 2.5$.

Since the purpose of the jj sample is to estimate the Z/γ^* background contributions, the Z/γ^* background contributions shown in the figures in this chapter are all derived from the jj sample data events. The details of the background estimation method will be discussed in Sec. 6.4.1.

6.2.4 Mass Reconstruction and Kinematic Likelihood Fit

Sharp resonances in the invariant masses of the scalars and the Higgs boson are powerful discriminants against background processes. Proper mass reconstructions of each scalar and the Higgs boson is thus crucial for this analysis. The energy resolution of reconstructed jet becomes worse as the p_{T} of the jet becomes low; lower the energy of the jet, the more it is susceptible to noise in the calorimeters and pile-up interactions. As the analysis looks for a four-body cascade decay from the Higgs boson with a mass of $m_H = 125$ GeV, each final state is expected to carry not much more than $\mathcal{O}(30)$ GeV of energy. This results in broadening of the resonance peak in the reconstructed mass distribution of $s \to bb$ decay. Fig. 6.3(b) shows the m_{bb} distributions of the bb sample. The solid line histograms show the m_{bb} distributions for the signal events, which have broad resonance peaks.

The resonance in the reconstructed mass distribution of $s \to \mu\mu$ decay, on the other hand, exhibits a sharp peak because of the good momentum resolution for muons in ATLAS. Fig. 6.3(a) shows sharp $m_{\mu\mu}$ distributions for the signal samples shown in solid line histograms. Broad m_{bb} distributions for signal samples leads to a broad resonance peak in the reconstructed Higgs boson mass distribution. Fig. 6.4(a) shows distributions of the invariant mass of the four fermion system $m_{bb\mu\mu}$ for the bb sample. The solid line histograms show the distributions of signal samples, which have broad resonance peaks.

To improve the resolution of the reconstructed Higgs mass, a kinematic likelihood fit (KLFit) procedure is performed. The procedure fits the energies of the jets to the constraint that the reconstructed mass of the s from the dijet and dimuon should be equal. For the bb sample, the signal jets are used in the KLFit, while the proxy jets are used for the jj sample in the KLFit.

The KLFit searches the energies of leading and subleading b-jet (E_{b_1} and E_{b_2}) that maximizes the likelihood for an event with measured leading and subleading b-jet energies ($\hat{E}_{b_1}$ and $\hat{E}_{b_2}$) with an invariant dimuon mass $m_{\mu\mu}$. The likelihood is defined as follows,

$$L = W(\hat{E}_{b_1}, E_{b_1}) \cdot W(\hat{E}_{b_2}, E_{b_2}) \cdot BW(m_{bb}, m_{\mu\mu}), \tag{6.2}$$

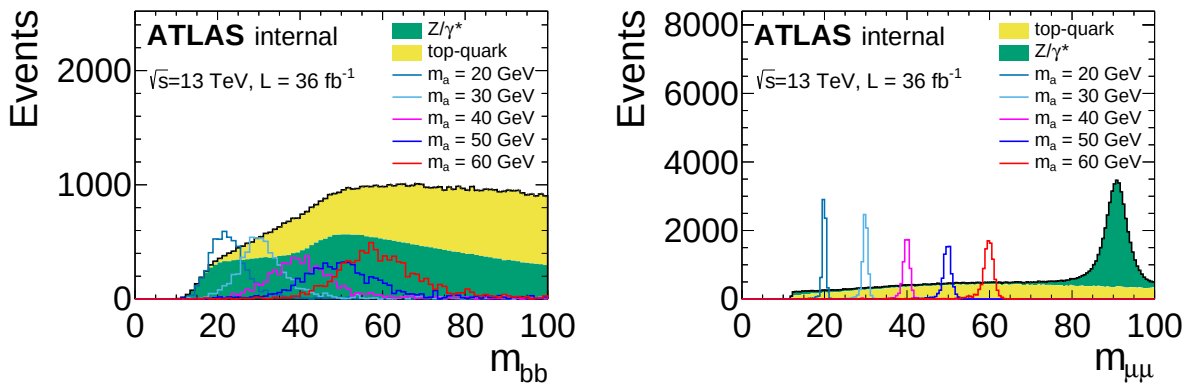
where m_{bb} is the recomputed invariant mass with the energy configuration that minimizes the likelihood, W is the transfer function (TF) of the b-jets, and the BW is a Breit-Wigner (BW) function. The transfer function $W(\hat{E}_{b_1}, E_{b_1})$ is a double gaussian probability density function such that, $\int W(\hat{E}, E)dE = 1$, and is parametrized as follows,

$$W(\hat{E}_b, E_b) = \frac{1}{\sqrt{2\pi}(p_2 + p_3 p_5)} \left(e^{-\frac{(\Delta E - p_1)^2}{2p_2^2}} + p_3 e^{-\frac{(\Delta E - p_4)^2}{2p_5^2}} \right), \quad \Delta E = \frac{\hat{E}_b - E_b}{\hat{E}_b}, \tag{6.3}$$

where the parameters p_i are functions of $\hat{E}$,

$$\begin{aligned} p_1 &= a_1 + b_1 \hat{E}, \\ p_2 &= a_2/\sqrt{\hat{E}} + b_2, \\ p_3 &= a_3 + b_3 \hat{E}, \\ p_4 &= a_4 + b_4 \hat{E}, \\ p_5 &= a_5 + b_5 \hat{E}, \end{aligned} \tag{6.4}$$

with parameters a_i and b_i are obtained through parametrized fits from MC simulated events. The BW



(a) (b)

Figure 6.3: (a) The m_{bb} distribution. (b) The $m_{\mu\mu}$ distribution. The signal events show broad m_{bb} distributions, while the $m_{\mu\mu}$ distributions show a sharp resonance as the detector has a good muon momentum resolution.

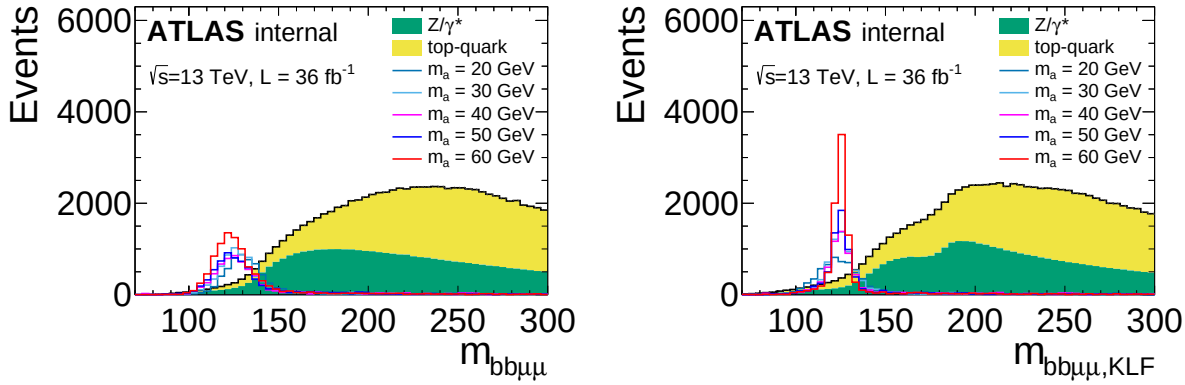
function is parametrized as follows,

$$BW(m_{bb}, m_{\mu\mu}) = \frac{k}{(m_{bb}^2 - m_{\mu\mu}^2)^2 + m_{\mu\mu}^2\Gamma^2}, \quad k = \frac{2\sqrt{2}m_{\mu\mu}\Gamma\sqrt{m_{\mu\mu}^2(m_{\mu\mu}^2+\Gamma^2)}}{\pi\sqrt{m_{\mu\mu}^2+\sqrt{m_{\mu\mu}^2(m_{\mu\mu}^2+\Gamma^2)}}}. \tag{6.5}$$

The KLFitter package [92] is used to implement the KLFit. By maximizing the likelihood, the KLFit is searching for a solution where the fitted energy values E_{b_1} and E_{b_2} are similar to the measured values $\hat{E}_{b_1}$ and $\hat{E}_{b_2}$, while ensuring the difference between the recomputed m_{bb} and the measured $m_{\mu\mu}$ is small. The fitted jet energies are then used to compute the invariant mass of the four-fermion system $m_{bb\mu\mu,\mathrm{KLF}}$ for the bb sample or $m_{jj\mu\mu,\mathrm{KLF}}$ for the jj sample. The recomputed $m_{bb\mu\mu}$ is called $m_{bb\mu\mu,\mathrm{KLF}}$. The subscript KLF denotes that the quantity uses the fitted jet energies. Fig. 6.4(b) shows the $m_{bb\mu\mu,\mathrm{KLF}}$ for the bb sample*. The signal histograms shown in solid lines exhibit much sharper structure after the KLFit.

The fit also outputs a likelihood value L, which signifies how well an event fits to the constraint. The log of the likelihood distribution of the fit is shown in Fig. 6.5 and the background distribution exhibits a clear bimodal shape where the lower half of the distribution can be removed without almost any loss of signal events. The $\ln(L)$ is used as a powerful discriminator against the background.

*The Z/γ^* contribution shown is actually the $m_{jj\mu\mu,\mathrm{KLF}}$ distribution from the jj sample as noted in Sec. 6.2.3.



(a) (b)

Figure 6.4: (a) The $m_{bb\mu\mu}$ distribution before the kinematic likelihood fit. (b) The $m_{bb\mu\mu,\mathrm{KLF}}$ distribution after the kinematic likelihood fit. The fit has resulted in much sharper signal peak, while the background distribution shape is largely unaltered. The sharpness of the peak will allow the signal-to-background ratio be improved.

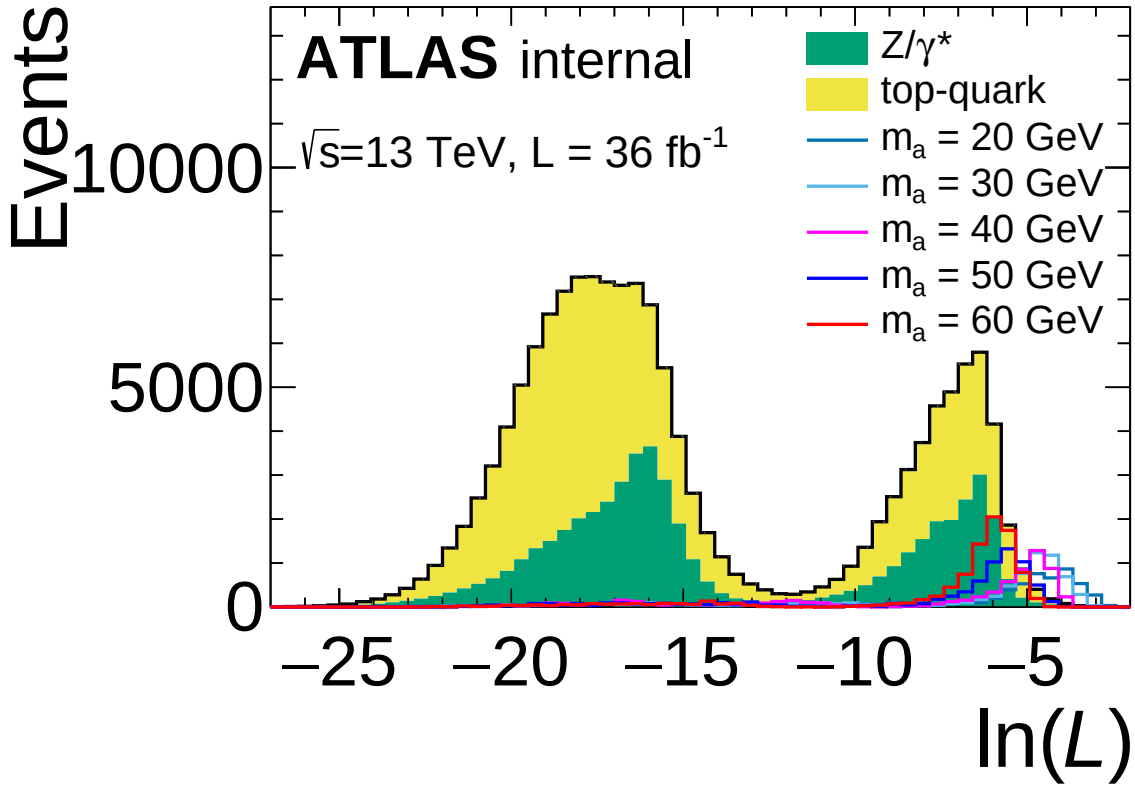


Figure 6.5: The L is the likelihood value of the kinematic fit result. Background events exhibit clear bimodal distribution. The lower end of the background events can be easily discarded without a significant loss of signal efficiency.

6.2.5 m_s Dependent bb and jj Samples

Both bb and jj samples are further split into five samples each by $|m_{\mu\mu} - m_s| < 5$ GeV for $m_s = 20$ GeV, 30 GeV, 40 GeV, 50 GeV, and 60 GeV. They are labeled by the m_s value in the subscript: bb_{m_s} or jj_{m_s} (e.g. bb_{20} or jj_{60}). For each bb_{m_s} or jj_{m_s} sample, a mass dependent $\ln(L)$ cut is applied. Selections of $\ln(L) > -6$, 5.5, 5, 6.5, 7 are required for samples with $m_s = 20$, 30, 40, 50, 60 GeV, respectively. As the signal process do not produce neutral stable particles, the $E_{\mathrm{T}}^{\mathrm{miss}}$ is expected to be low, while the top-quark process contains neutrinos in the final states producing sizable $E_{\mathrm{T}}^{\mathrm{miss}}$. Selections of $E_{\mathrm{T}}^{\mathrm{miss}} < 60$, 50, 50, 30, 30 GeV are required for samples with $m_s = 20$, 30, 40, 50, 60 GeV, respectively.

6.2.6 Summary

There are total of 10 samples defined: $bb_{20,30,40,50,60}$ samples and $jj_{20,30,40,50,60}$ samples. The details of selections applied to define each sample is summarized in Table 6.1. The jj_{m_s} samples are used to estimate Z/γ^* contributions in bb_{m_s} samples.

6.3 Signal Extraction Strategy

The feature of sharp $m_{\mu\mu}$ distribution is used to search for signal events; the $m_{\mu\mu}$ distribution will be fitted to the data. As the search focuses on a non-SM decay of the Higgs boson, a requirement of $|m_{bb\mu\mu,\mathrm{KLF}} - m_H| < 15$ GeV is applied to all events.

Fig. 6.6 shows the final fit region with proper background estimation applied. The background estimation will be discussed in Sec. 6.4. The signal is scaled down to a cross section corresponding to $\mathcal{B}(H \to ss) = 30\%$ for plotting purpose. High signal-to-background ratios are observed for all mass points studied. The discontinuities in the distribution are due to m_s dependent selections. For the final fit, much finer binnings are used. The binning used for the fit will be discussed in Sec. 6.5.

6.4 Background Estimation

6.4.1 Z/γ^* Background Estimation

Several control regions are used to estimate Z/γ^* background. The bb and jj samples are each split into two subsamples. The events in bb sample satisfying $|m_{bb\mu\mu,\mathrm{KLF}} - m_H| < 15$ GeV are called the bb^{H} sample. The H indicates that a Higgs mass window selections are applied to this sample. The events in bb sample satisfying $|m_{bb\mu\mu,\mathrm{KLF}} - m_H| > 15$ GeV and $|m_{bb\mu\mu,\mathrm{KLF}} - m_H| < 45$ GeV are called the bb^{SB} sample. The

Table 6.1: Summary of selection critera and sample definitions. Trigger and data quality requirements are not listed but also applied to all events.

Preselection	$p_{\mathrm{T}}^{\mu_1} > 27$ GeV and $p_{\mathrm{T}}^{\mu_2} > 7$ GeV Veto events with $n_\ell > 2$ with $p_{\mathrm{T}}^{\ell 3} > 7$ GeV Opposite-charge leptons $n_b \geq 2$ $p_{\mathrm{T}}^{b_{1,2}} \geq 20$ GeV $m_{\mu\mu} > 12$ GeV $m_{\mu\mu} < 80$ GeV $n_j \geq 2$
Separation of bb and jj samples	bb sample : $n_b \geq 2$ jj sample : $n_b \leq 1$
Separation by m_s	bb_{20} and jj_{20} sample : $\lvert m_{\mu\mu} - 20 \text{ GeV}\rvert < 5$ GeV bb_{30} and jj_{30} sample : $\lvert m_{\mu\mu} - 30 \text{ GeV}\rvert < 5$ GeV bb_{40} and jj_{40} sample : $\lvert m_{\mu\mu} - 40 \text{ GeV}\rvert < 5$ GeV bb_{50} and jj_{50} sample : $\lvert m_{\mu\mu} - 50 \text{ GeV}\rvert < 5$ GeV bb_{60} and jj_{60} sample : $\lvert m_{\mu\mu} - 60 \text{ GeV}\rvert < 5$ GeV
$m_{\mu\mu}$ dependent likelihood selection	bb_{20} and jj_{20} sample : $\ln(L) > -6$ bb_{30} and jj_{30} sample : $\ln(L) > -5.5$ bb_{40} and jj_{40} sample : $\ln(L) > -5$ bb_{50} and jj_{50} sample : $\ln(L) > -6.5$ bb_{60} and jj_{60} sample : $\ln(L) > -7$
$m_{\mu\mu}$ dependent $E_{\mathrm{T}}^{\mathrm{miss}}$ selection	bb_{20} and jj_{20} sample : $E_{\mathrm{T}}^{\mathrm{miss}} < 60$ GeV bb_{30} and jj_{30} sample : $E_{\mathrm{T}}^{\mathrm{miss}} < 50$ GeV bb_{40} and jj_{40} sample : $E_{\mathrm{T}}^{\mathrm{miss}} < 50$ GeV bb_{50} and jj_{50} sample : $E_{\mathrm{T}}^{\mathrm{miss}} < 40$ GeV bb_{60} and jj_{60} sample : $E_{\mathrm{T}}^{\mathrm{miss}} < 30$ GeV

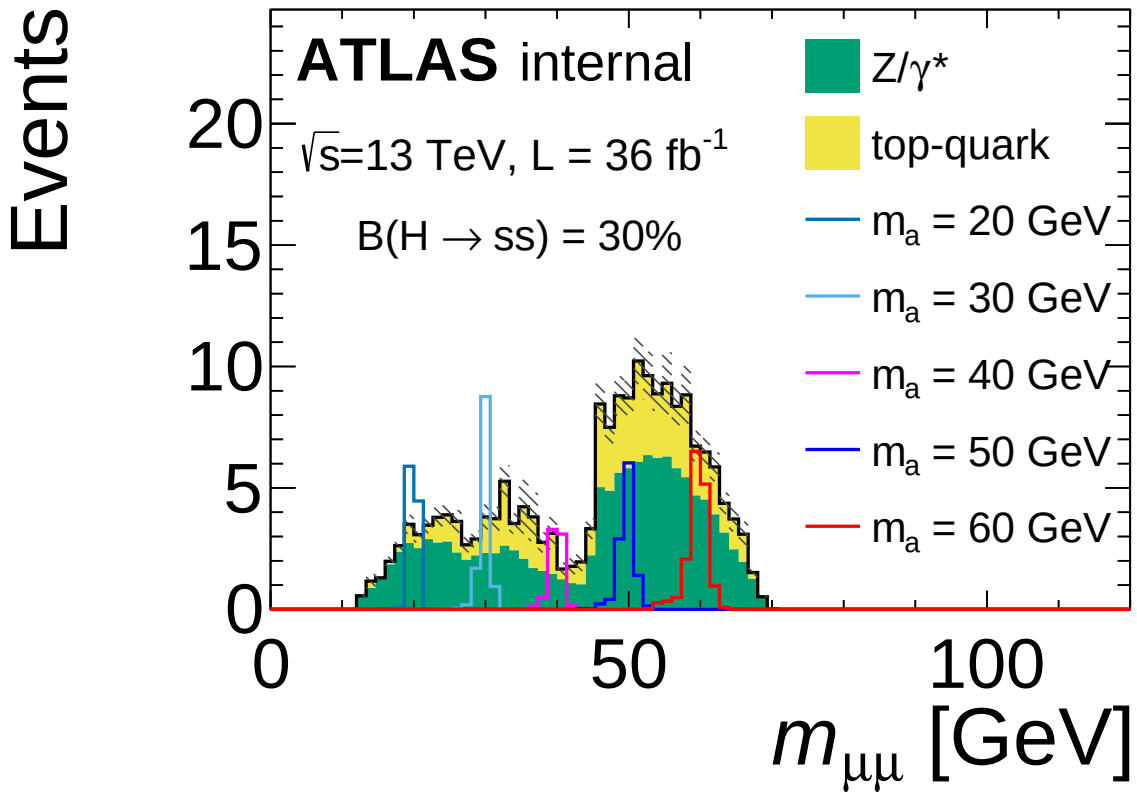


Figure 6.6: The final fit region $m_{\mu\mu}$ distribution. For the final fit, the binning will be made much finer.

SB indicates side-band regions around the Higgs boson mass window. Analogous selections are made to the jj sample using the $m_{jj\mu\mu,\mathrm{KLF}}$ variable instead and two subsamples jj^{H} and jj^{SB} are formed. The bb^{H} and jj^{H} are referred to as the Higgs mass window samples, and bb^{SB} and jj^{SB} are referred to as the side-band region samples.

Fig. 6.7 shows the illustration of the different regions defined in previous paragraph. There are four colored regions shown in Fig. 6.7 (n.b. the bb^{SB} and jj^{SB} are repeated twice). The y-axis represents the number of b-jets. The top rows are the events that pass $n_b \geq 2$ and the botom rows are the events that pass $n_b \leq 1$. The x-axis represents the invariant mass of the four-fermion system. Depending on the corresponding number of b-jets, the x-axis variable is either $m_{bb\mu\mu,\mathrm{KLF}}$ or $m_{jj\mu\mu,\mathrm{KLF}}$. The value of m_H is marked with a 30 GeV window around it to illustrate the window cut applied on $m_{bb\mu\mu,\mathrm{KLF}}$ or $m_{jj\mu\mu,\mathrm{KLF}}$; As explained in previous paragraph, the events in the bb^{H} (jj^{H}) sample are required to pass the $|m_{bb\mu\mu,\mathrm{KLF}} - m_H| < 15$ GeV ($|m_{jj\mu\mu,\mathrm{KLF}} - m_H| < 15$ GeV) criterion. The final fit is performed on events in the bb^{H} sample.

It is important to note that both jj^{H} and jj^{SB} samples are predominantly Z/γ^* events with high purity. Suppose the shape of the $m_{\mu\mu}$ distribution of Z/γ^* events in the jj^{H} sample is the same as the $m_{\mu\mu}$ distribution of Z/γ^* events in the bb^{H} sample. Then, combined with the fact that the jj^{H} sample purely consists of Z/γ^* background, the $m_{\mu\mu}$ shape observed the jj^{H} sample can be directly taken to model the $m_{\mu\mu}$ distribution in the bb^{H} sample. This determines the shape of the $m_{\mu\mu}$ distribution of Z/γ^* background in the bb^{H}.

The validity of this assumption is checked with the MC samples and is confirmed within the statistical precision of the MC samples. Fig. 6.8 shows the check using MC sample with a caveat that the uniform $\ln(L) > -11$ and $E_{\mathrm{T}}^{\mathrm{miss}} < 50$ GeV selections were applied across all $m_{\mu\mu}$, instead of the mass dependent selections. The top panel shows the MC events in the $n_b \leq 1$ region, whereas the bottom panel shows the MC events in the fit region. The ratio R between the number of MC events across $m_{\mu\mu}$ shows a flat distribution confirming within the statistical precision of the MC sample that the two regions have the same $m_{\mu\mu}$ distribution. When the mass dependent selections are applied, the thresholds on the $E_{\mathrm{T}}^{\mathrm{miss}}$ and $\ln(L)$ selections are applied the same for both bb and jj sample, and the conclusion is not expected to change as the $E_{\mathrm{T}}^{\mathrm{miss}}$ and $\ln(L)$ distributions also show good shape agreements between the two samples [93].

The events in side-band region samples are kinematically similar to the events in the Higgs mass window samples, as the two only differ by few GeV in $m_{bb\mu\mu,\mathrm{KLF}}$ or $m_{jj\mu\mu,\mathrm{KLF}}$. Therefore, the events in the side-band regions serve as a control sample for estimating properties of the Higgs mass window samples. From the side-band regions, an extrapolation factor α is measured. The α is defined as the ratio of Z/γ^* contributions

in $bb^{\rm SB}$ to $jj^{\rm SB}$,

$$\alpha \equiv \frac{N^{Z/\gamma^*}_{bb^{\rm SB}}}{N^{Z/\gamma^*}_{jj^{\rm SB}}}, \tag{6.6}$$

where N^{Z/γ^*}_R is the size of the Z/γ^* background contribution in sample R. Since the events in side-band region samples are similar to the events in Higgs mass window samples, the following relation also holds,

$$\alpha \simeq \frac{N^{Z/\gamma^*}_{bb^{\rm H}}}{N^{Z/\gamma^*}_{jj^{\rm H}}}. \tag{6.7}$$

The validity of the relation has been checked using MC samples and has been confirmed within the statistical precision of the MC samples. Therefore, the measured α can be used to estimate the Z/γ^* background contribution in the $bb^{\rm H}$ sample,

$$N^{Z/\gamma^*}_{bb^{\rm H}} \simeq N^{Z/\gamma^*}_{jj^{\rm H}} \cdot \alpha. \tag{6.8}$$

This determines the overall normalization of the expected Z/γ^* background contributions in the fit region.

Fig. 6.9 shows the $m_{bb\mu\mu,\rm KLF}$ or $m_{jj\mu\mu,\rm KLF}$ distributions for the four regions. The Fig. 6.9(c) shows the data events in the $jj^{\rm SB}$ with tiny top-quark background contributions that is almost invisble; the top-quark contribution is ignored. The template from Fig. 6.9(c) is carried over to the region $bb^{\rm SB}$ shown in Fig. 6.9(a) and scaled down to fit the data. Once a good fit is produced, α is measured by the ratio between the Z/γ^* contribution in Fig. 6.9(a) and Fig. 6.9(c) The same α is then applied to the distribution shown in Fig. 6.9(d) and the result is plotted in Fig. 6.9(b) as the Z/γ^* contribution. For the final fit procedure, the regions shown in Fig. 6.9(a) and Fig. 6.9(b) will be fitted simultaneously.

The Fig. 6.10(a) shows the $m_{\mu\mu}$ distribution in the $bb^{\rm SB}$ region after the α factor is applied. Despite mass dependent cuts applied in different $m_{\mu\mu}$ regions, the modeling of the Z/γ^* distribution shows very good agreement with the data adding confidence to the method. It is important to note that only the shape should be compared here, as the normalization is in a good agreement by construct due to the fitting procedure. To further validate the procedure, a high side-band control region is used to verify the validity of the extrapolation factor. The high side-band region is defined to be the region above $m_{bb\mu\mu,\rm KLF} > 170$ GeV for the bb sample. The same α measured from the fit to the data shown in Fig. 6.9(a) is applied to data events in the jj sample with $m_{jj\mu\mu,\rm KLF} > 170$ GeV to validate that the modeling of various kinematics. Fig. 6.10(b)–6.10(e) shows the several key distributions in the analysis and all distributions show good agreements in both the normalization and the shape.

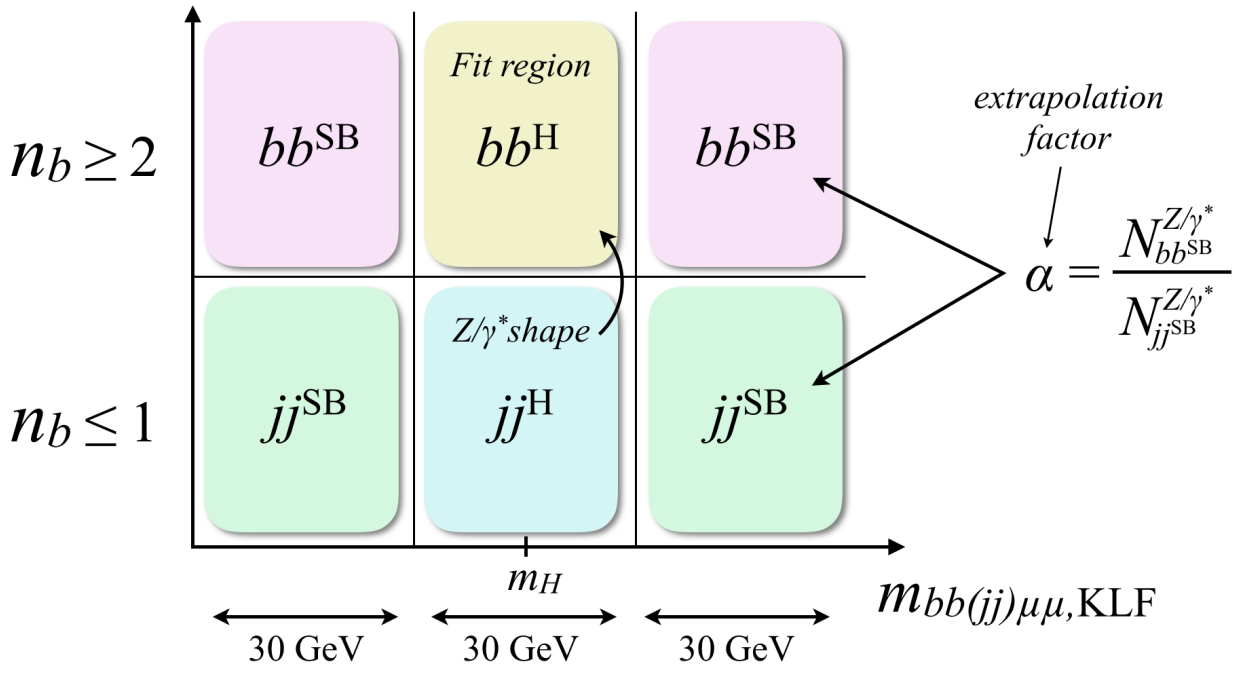


Figure 6.7: Illustration of different regions. The regions denoted by superscript $^{\mathrm{SB}}$ are the side-band regions, where the extrapolation factor α is measured.

6.4.2 Top-quark Background Estimation

The top-quark background is estimated using Method B described in Sec. 5.4.1. The control region is motivated by the fact that the leptonically decaying top-quark process produces neutrinos, so the control region is defined to be the same as the fit region, except that the selections on $E_{\mathrm{T}}^{\mathrm{miss}}$ are inverted. The event sample in the control region has high top-quark process purity, but some residual Z/γ^* background contamination exists. In order to obtain the observed number of top-quark contribution in the control region, the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution is fitted. The template for top-quark contribution is obtained from the MC sample, and the Z/γ^* background $E_{\mathrm{T}}^{\mathrm{miss}}$ template is obtained from the jj sample with event selections identical to the top control region definition.

Fig. 6.11 shows the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in the top control region. The shape of the top-quark and Z/γ^* contributions show difference in shape. By fitting both templates from each background to the data events at the same time, one can obtain the relative contributions to each. Before the fit, the total size of the top-quark background estimated in the control region is the $\bar{B}_{\mathrm{CR}}$ of Eq. (5.5). Once a good fit is obtained, the size of the top-quark background can be measured. The measured quantity will be the $\bar{N}_{\mathrm{CR}}$ of Eq. (5.5). Using the ratio $\beta \equiv \bar{N}_{\mathrm{CR}}/\bar{B}_{\mathrm{CR}}$, the final estimation of the top-quark background in each $m_{\mu\mu}$ bin of the fit region is obtained by multiplying the β to the number of expected top-quark background in each $m_{\mu\mu}$ bin

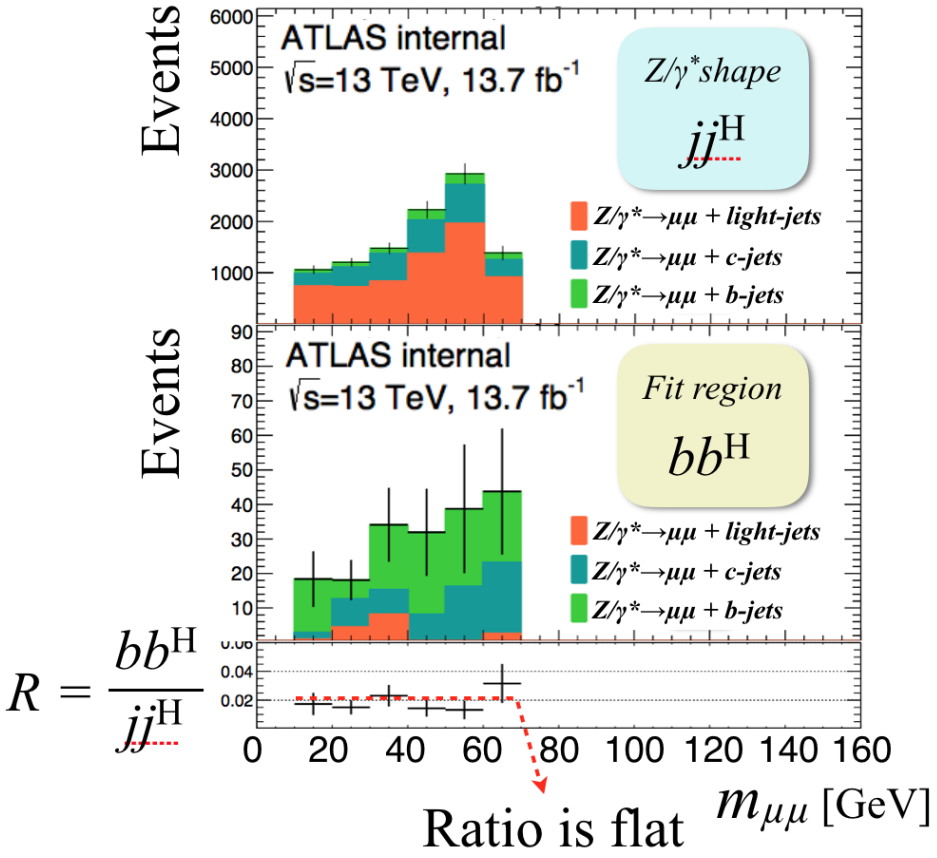


Figure 6.8: The $m_{\mu\mu}$ distributions in the jj^{H} (top) and bb^{H} (bottom) regions with Z/γ^* MC samples. The MC samples are separated by flavor compositions. The ratio of the two samples are shown in the bottom panel. The ratio shows a flat distribution, indicating that the $m_{\mu\mu}$ shape in the two regions are the same [93].

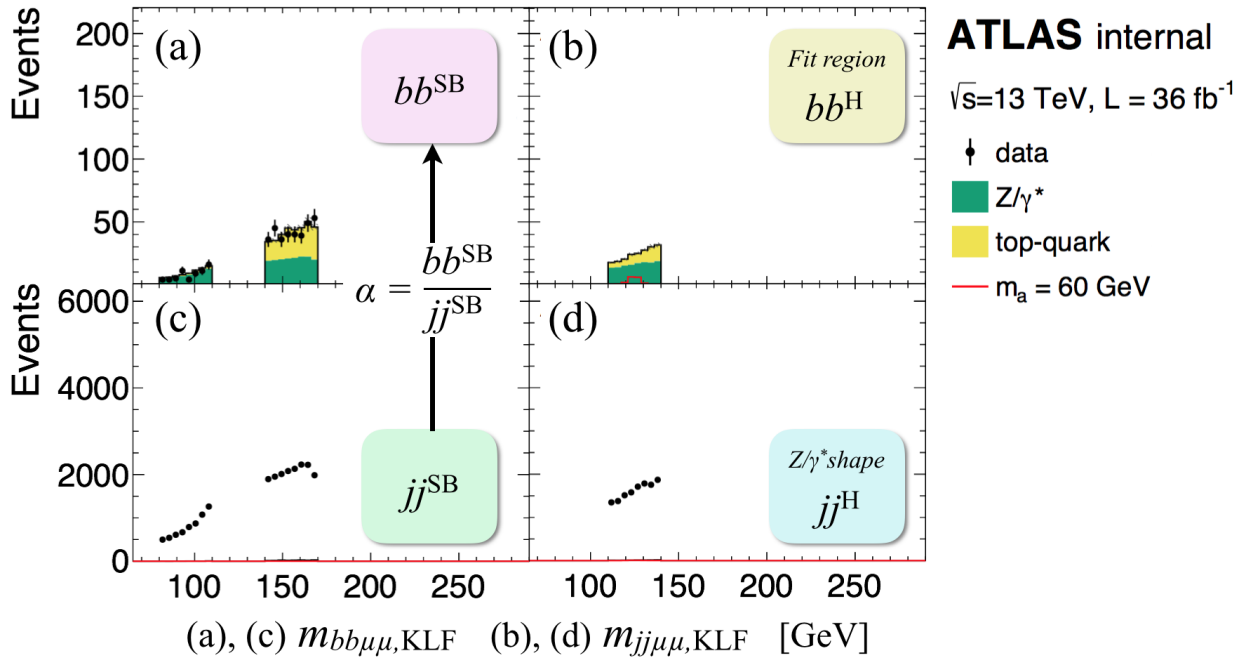


Figure 6.9: Illustration of background estimation. The data in region (b) is used as a template to fit to the data in (a) along with the top-quark background MC samples. After the fit, the α is computed by taking the ratio of the Z/γ^* contributions from (a) and (c). The obtained α is applied to the data events in (d) to model the expected Z/γ^* contributions in region (a).

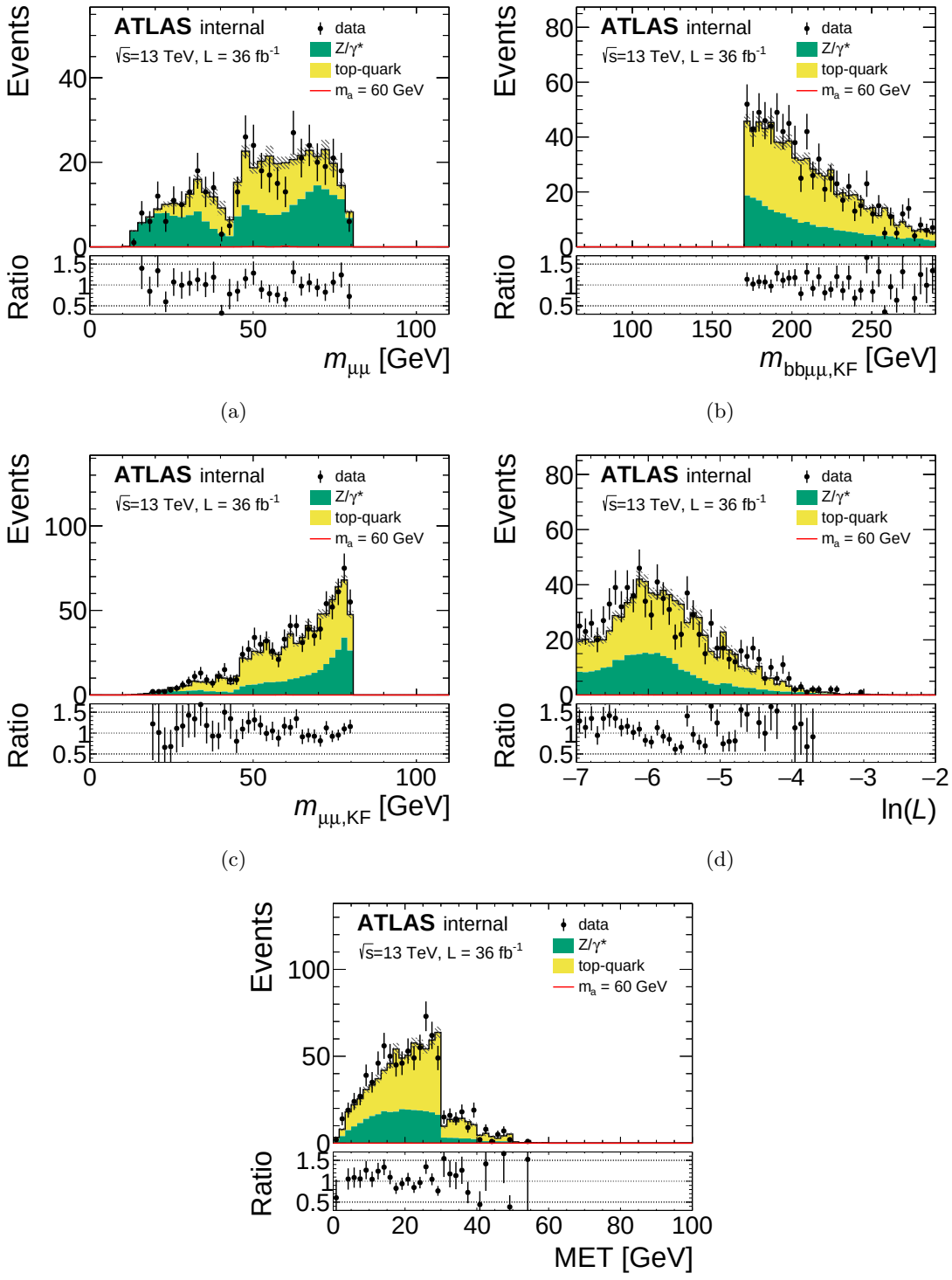
by the MC sample B_{SR}.

6.5 Fit Procedure and Expected Sensitivity

There are three regions simultaneously being fitted. The first region is the SB region shown in Fig. 6.9(a). The second region is the top-quark control region shown in Fig. 6.11. The third region is the signal region appropriate for each bb_{m_s} samples. The Fig. 6.12 shows the fit regions for the five mass points.

The likelihood of the fit is defined the same as discussed in Sec. 5.6. The Fig. 6.9(a) and Fig. 6.11 forms the control regions, and there are two β parameters being probed where each normalizes the size of Z/γ^* and top-quark backgrounds. The respective β parameters then defines expected background contribution in the Poisson function $f(N|\mu S + \beta_{Z/\gamma^*} B_{Z/\gamma^*} + \beta_{\text{top}} B_{\text{top}})$ of the signal region. The implementation of the fit is done through HistFitter [94] for fitting, profiling, and establishing upper limts on μ using the CL_s method [95–97].

The expected sensitivity is shown in Fig. 6.13. The limit is expressed in terms of the $\mathcal{B}(H \to ss)$. The



(a) (b) (c) (d) (e)

Figure 6.10: Several validation plots of the Z/γ^* background estimation method. (a) shows the $m_{\mu\mu}$ distribution in the side-band region to check that the shape is in good agreement with data. (b)-(e) shows four different key variables in the validation region to check both the shape and the normalization of the estimated Z/γ^* background against the data distributions. In all cases, the background is well modeled.

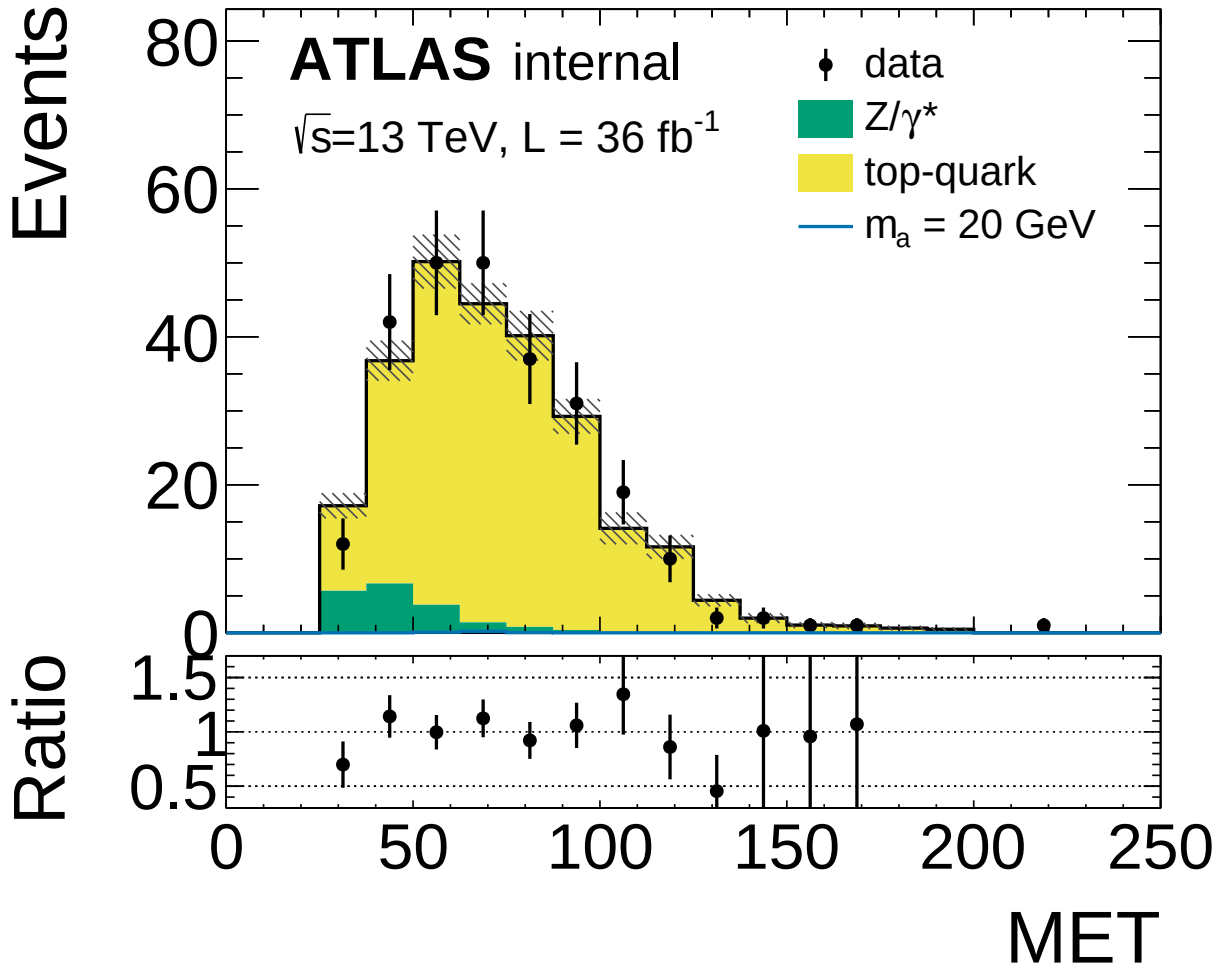


Figure 6.11: The $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in the top control region used to estimate the top-quark background process. The small Z/γ^* contribution is modeled by the corresponding jj sample events. The $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution shown here is used during the fit to profile top-quark background contribution in the signal fit region.

search is expected to exclude at 95% confidence level a branching fraction of,

$$\mathcal{B}(H \to ss) = 15 - 30\%, \tag{6.9}$$

for the mass range $m_s = 20 - 60$ GeV, assuming type-III 2HDM+S model with $\tan\beta = 2$ which results in $\mathcal{B}(ss \to bb\mu\mu) = 1.6 \times 10^{-3}$. This number can be reinterpreted as the exclusion limit on,

$$\mathcal{B}(H \to ss \to bb\mu\mu) = 2.4 - 4.8 \times 10^{-4}. \tag{6.10}$$

The expected sensitivity of the presented direct search constrains better than the indirect exclusion limit from combined Higgs couplings measurement shown in Ref. [22].

6.6 Discussion on Systematic Uncertainties

The systematic uncertainties are not included in the results shown in Sec. 6.5 mainly for the reason that the statistical uncertainty in the fit region is the dominant source of uncertainty. Fig. 6.12 shows that the each bin generally expects about one or two events, so the statistical uncertainty in each bin will be $\mathcal{O}(50-100\%)$. Most of the uncertainties from the PID are on the order of $\mathcal{O}(1-10\%)$. The uncertainty associated with the background estimation method for the Z/γ^* background can be assessed from the check done in Fig. 6.8 using the MC samples. It shows around $\mathcal{O}(20-40\%)$ differences, but with higher statistics MC sample, the uncertainty is expected to shrink. The top-quark MC modeling tends to add systematic uncertainties of order $\mathcal{O}(20\%)$, similar to the VBF analysis in Chapter 5. Overall, the systematic uncertainty is expected to be smaller than that of the statistical uncertainty.

Despite the fact that the systematic uncertainty is expected to be low, a check is performed by adding few mock-up systematics to assess the effect. An uncorrelated systematics of 20% and a corrlecated systematics of 10% are added to the top-quark background An uncorrelated systematics of 40% is added to the Z/γ^* background as a mock-up of the uncertainty on the background estimation method. A correlated systematics of 15% is added to the signal process. As expected, with the additional uncertainties, exclusion limits on each mass point degraded by only $1-2\%$p. A proper study with the relevant systematics will be left for a future study.

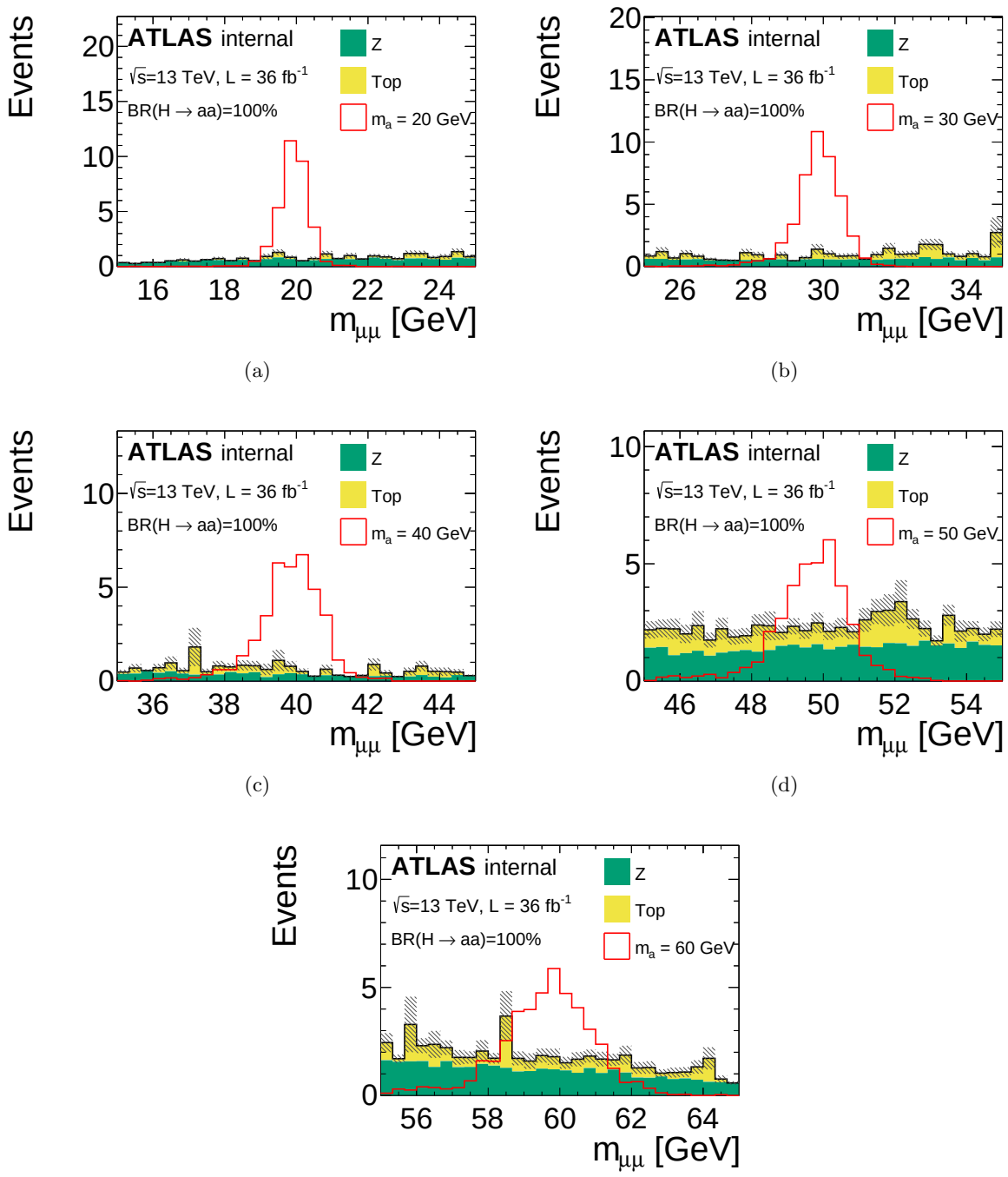
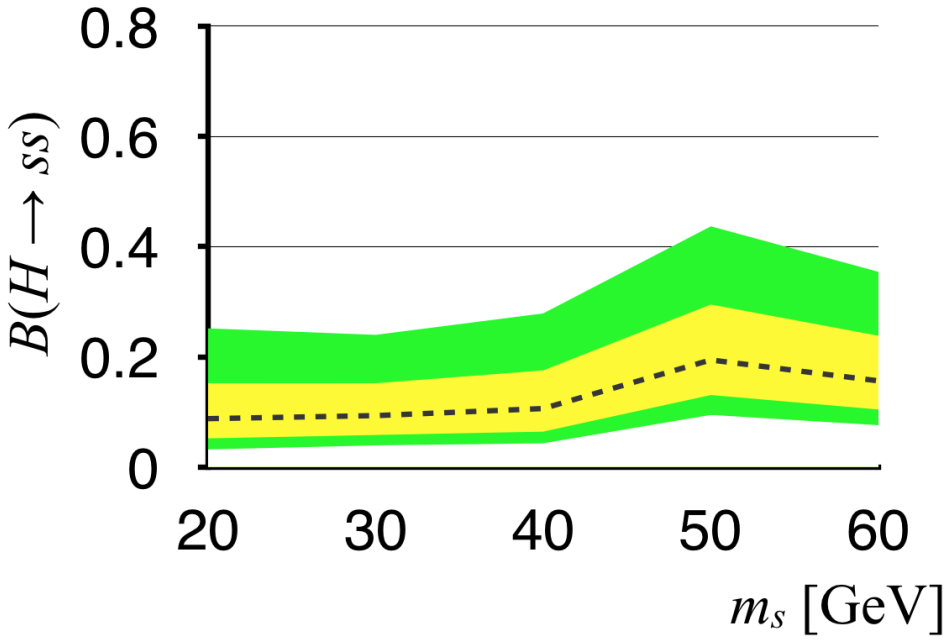


Figure 6.12: The $m_{\mu\mu}$ distributions used in the final fit.



(a)

Figure 6.13: Expected 95% confidence limit exclusion limit on branching fraction of $\mathcal{B}(H \to ss) = 15 - 30\%$ for the mass range $m_s = 20 - 60$ GeV, assuming type-III 2HDM+S model with $\tan\beta = 4$ which results in $\mathcal{B}(ss \to bb\mu\mu) = 1.6 \times 10^{-3}$.

Chapter 7

Discussions and Suggestions for Future Work

7.1 Higgs Couplings Measurement

In Chapter 5, the measurement of the ratio of the signal strengths $\mu_{\mathrm{VBF}}/\mu_{\mathrm{ggF}}$ is presented. The same global fit from the analysis can be used to profile the μ_{VBF} and μ_{ggF} individually. The Fig. 7.1 shows the likelihood scan in terms of the two parameters. The result of the best-fit signal strengths are,

$$\begin{aligned} \mu_{\mathrm{ggF}} &= 1.02 \quad \pm 0.19 \quad {}^{+0.22}_{-0.18} \quad = 1.02 \quad {}^{+0.29}_{-0.26} \\ \mu_{\mathrm{VBF}} &= 1.27 \quad {}^{+0.44}_{-0.40} \quad {}^{+0.30}_{-0.21} \quad = 1.27 \quad {}^{+0.53}_{-0.45}. \\ &\qquad\quad (\mathrm{stat}) \quad (\mathrm{syst}) \end{aligned} \tag{7.1}$$

The μ_{ggF} shows a precision of around 30%, while the precision on the μ_{VBF} is around 50% with a slight excess, which is statistically insignificant. The results show that the Higgs boson signal strength measurement is in agreement with the SM prediction. The systematic uncertainty component contributing to the total uncertainty is subdominant but the statistical uncertainty contributes at a similar level.

The measured signal strength μ's can be directly related to the strengths of the Higgs couplings. According to the narrow width approximation, the cross section of a process shown in Fig. 7.2 can be expressed as,

$$\sigma \propto \frac{g_p^2 g_d^2}{\Gamma_H}, \tag{7.2}$$

where the g_p is the coupling strength of the production vertex, the g_d is the coupling strength of the decay vertex, and Γ_H is the total width of the Higgs boson. As the signal strengths are the ratio of the observed to expected number of signal events, the μ can be represented as,

$$\mu = \frac{\hat{\sigma}}{\sigma} = \frac{\kappa_p^2 \kappa_d^2}{\kappa_H^2} \tag{7.3}$$

where $\kappa_{p,d} = \hat{g}_{p,d}^2/g_{p,d}^2$ and $\kappa_H^2 = \hat{\Gamma}_H/\Gamma_H$, with hats on variables representing the observed values from

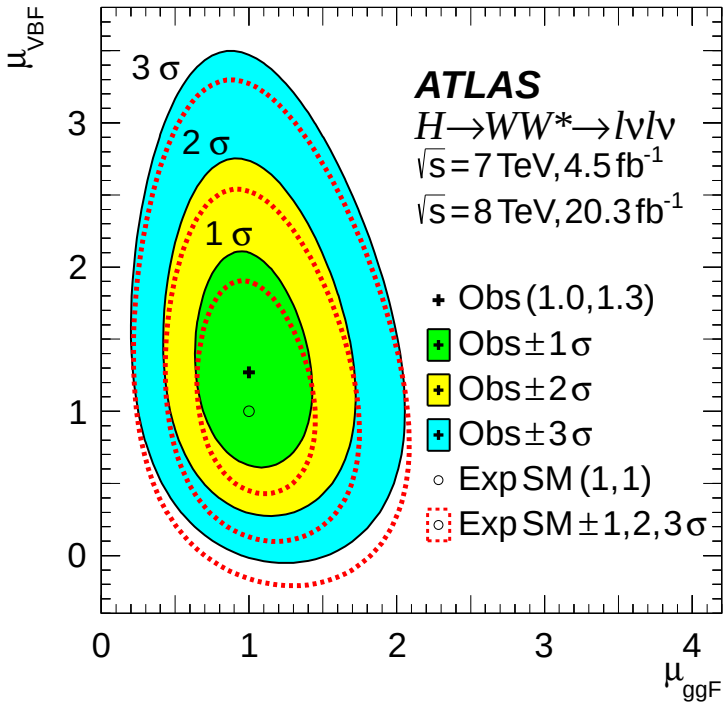


Figure 7.1: The figure and the caption is reproduced from Ref. [18]. Likelihood scan as a function of μ_{ggF} and μ_{VBF}. The best-fit observed (expected SM) value is represented by the cross symbol (open circle) and its one, two, and three standard deviation contours are shown by solid lines surrounding the filled areas (dotted lines). The x- and y-axis scales are the same to visually highlight the relative sensitivity.

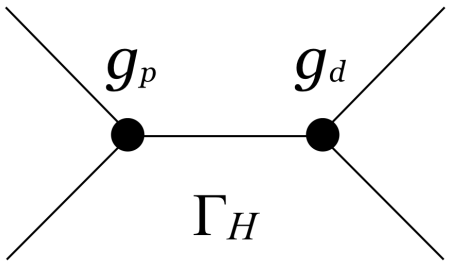


Figure 7.2: Feynman diagram of Higgs production and decay. The production vertex has a coupling strength of g_P, and the decay vertex has a coupling strength of g_D. The total width of the Higgs boson is Γ_H.

data. This parametrization assumes that the Higgs boson does not decay to BSM particles. As discussed in the Sec. 1.7, the total width of a particle is proportional to the sum of all couplings squared. Hence the modification to the Γ_H can be written as,

$$\kappa_H^2 = \sum_i \mathcal{B}_i \kappa_i^2 \tag{7.4}$$

where $\mathcal{B}_i$'s are the branching ratios. Despite the Higgs couplings to different final states have different coupling strengths, one hypothesizes a common coupling modifier κ_F for fermionic couplings and another modifier κ_V for bosonic couplings. In this parametrization, the signal strengths becomes,

$$\mu_{\mathrm{VBF}} \propto \frac{\kappa_V^4}{\mathcal{B}_F \kappa_F^2 + \mathcal{B}_V \kappa_V^2}, \quad \mu_{\mathrm{ggF}} \propto \frac{\kappa_V^2 \kappa_F^2}{\mathcal{B}_F \kappa_F^2 + \mathcal{B}_V \kappa_V^2}, \tag{7.5}$$

where $\mathcal{B}_F$ represents the total branching ratios to final states that is mediated by the fermionic couplings of the Higgs boson, and vice versa for $\mathcal{B}_V$. For the Higgs boson with the mass of $m_H = 125.36$ GeV, the branching ratios are $\mathcal{B}_F \approx 75\%$ and $\mathcal{B}_V \approx 25\%$. For large κ_F values, the denominators in Eq. (7.5) approximates to $\mathcal{B}_F \kappa_F^2 + \mathcal{B}_V \kappa_V^2 \approx \kappa_F^2$. This implies that for large κ_F values the dependence on κ_F in μ_{ggF} roughly cancels out. The μ_{VBF} on the other hand becomes,

$$\mu_{\mathrm{VBF}} \propto \kappa_V^4 / \kappa_F^2, \tag{7.6}$$

and does not exhibit canceling behavior. If κ_F is large, the μ_{VBF} is small. Exclusion on $\mu_{\mathrm{VBF}} = 0$ thus puts limit on how big κ_F can be, while the μ_{ggF} only helps constraining κ_V when $\kappa_F \gg \kappa_V$. This implies that *the most of the sensitivity for excluding large κ_F values in WW final state channel comes from measuring VBF process signal strength.*

Using this parametrization the likelihood can be scanned in terms of κ_F vs. κ_V. The scan is shown in Fig. 7.3. The dashed lines show the expected contours and the solid colors show the observed contours for one, two and three standard deviations. The best-fit value is marked by the cross and they are,

$$\begin{aligned} \kappa_F &= 0.93 \;{}^{+0.24}_{-0.18}\;{}^{+0.21}_{-0.14} = 0.93 \;{}^{+0.32}_{-0.23} \\ \kappa_V &= 1.04 \;{}^{+0.07}_{-0.08}\;{}^{+0.07}_{-0.08} = 1.04 \pm 0.11. \\ &\qquad\quad (\mathrm{stat}) \quad (\mathrm{syst}) \end{aligned} \tag{7.7}$$

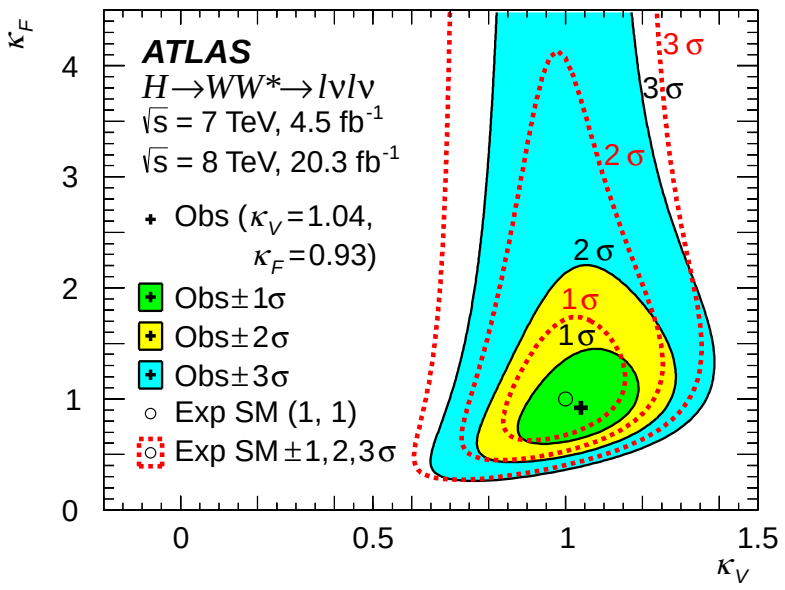


Figure 7.3: The figure and the caption is reproduced from Ref. [18]. Likelihood scan as a function of κ_V and κ_F. The best-fit observed (expected SM) value is represented by the cross symbol (open circle) and its one, two, and three standard deviation contours are shown by solid lines surrounding the filled areas (dotted lines). Note that the y-axis spans a wider range than the x-axis.

7.2 Combinations with other Higgs Boson Measurement

The signal strengths for different Higgs boson channels have been measured in ATLAS and CMS. The result presented in the thesis is included in the combination of all Higgs boson measurements from ATLAS [98]. The Fig. 7.4 shows the result of overall Higgs boson signal strength for the five major decay channels. The $H \to WW^*$ channel provides the most precise signal strength measurements with uncertainty on the signal strength $\Delta\mu = \mathcal{O}(20\%)$. The $H \to \gamma\gamma$ and $H \to ZZ^*$ channels, the golden channels for the discovery, in fact, shows less precision than the $H \to WW^*$.

The importance of $H \to WW^*$ channel relative to other decay channels becomes even more clearer when the signal strength measurements are reinterpreted as constraints on κ_F and κ_V. The Fig. 7.5 shows the combined κ_F and κ_V measurement in black contours. The individual channel contributions are shown in different contours. The gray color contour shows the $H \to WW^*$ contribution, which has the smallest contour among the five decay channels. It is the only channel that significantly constrains both the κ_F and κ_V. It is worth noting that the diagonal shaped constraint from $H \to \gamma\gamma$ determines that the sign of the κ_F and κ_V as positive. This is because the signal strengths of the $H \to \gamma\gamma$ process contains interference terms.

The Compact Muon Solenoid (CMS) experiment has also conducted similar searches [100]. The CMS result of the κ_F and κ_V measurement is shown in Fig. 7.6. The results for the five individiual channels are

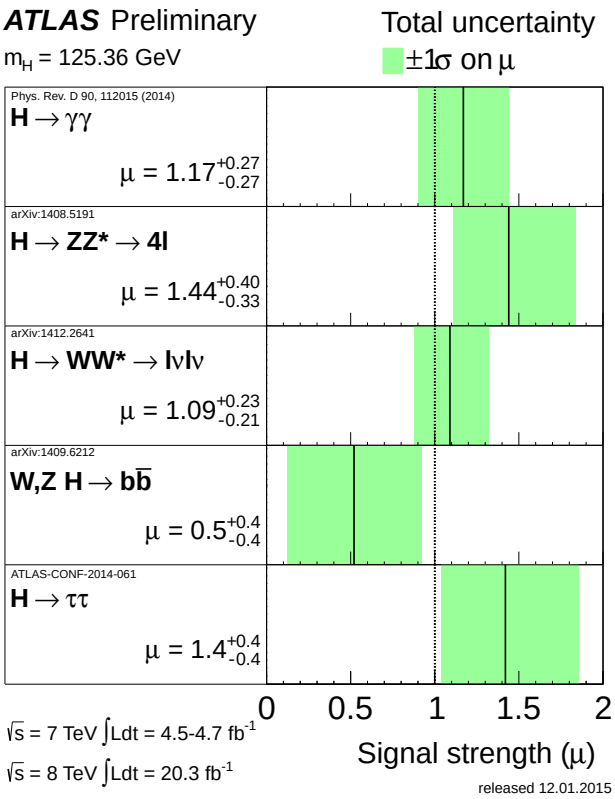


Figure 7.4: Summary plot showing the signal strength measurements of the five major decay channels from ATLAS: $H \to \gamma\gamma$, $H \to ZZ^*$, $H \to WW^*$, $H \to b\bar{b}$, and $H \to \tau\tau$. The most precise signal strength measurement is from $H \to WW^*$ [99].

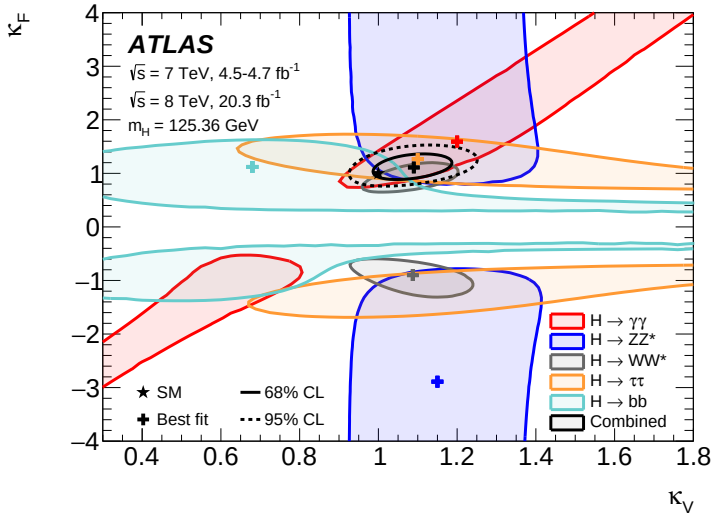


Figure 7.5: Constraint on κ_F vs. κ_V from the five major decay channels from ATLAS. The combined κ_F vs. κ_V constraint are shown along with the breakdown of each decay channel. The best constraint on both κ_F and κ_V is from $H \to WW^*$ [98].

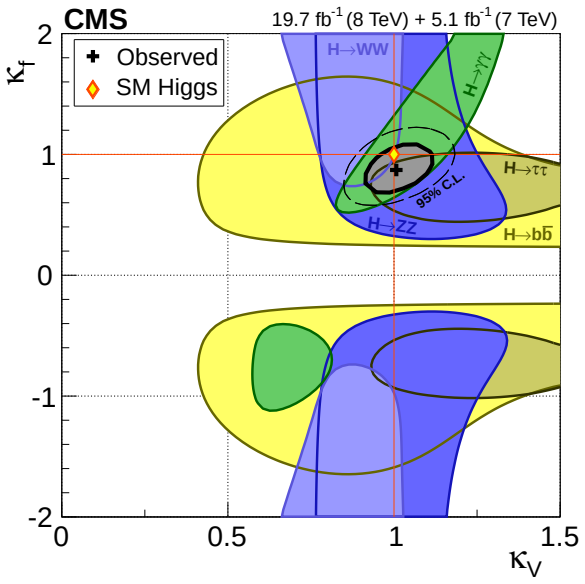


Figure 7.6: Constraint on κ_F vs. κ_V from the five major decay channels from CMS. The combined κ_F vs. κ_V constraint are shown along with the breakdown of each decay channel [100].

shown in colored swaths and the light blue area indicates the constraint from $H \to WW^*$. It can be seen that the high κ_F area is not constrained by the $H \to WW^*$ results from CMS. The difference can be understood from the lack of sensitivity in VBF production process in the $H \to WW^*$ analysis from CMS [101]. The analysis from CMS does not adopt BDT algorithm for an improved sensitivity and the result does not show a significant excess.

The grand combination results of κ_F and κ_V of all Higgs boson measurements of ATLAS and CMS is shown in Fig. 7.7 [22]. In this fit the κ_F and κ_V values are constrained to be positive. This shows the state of the art measurement of the κ_F and κ_V of the Higgs boson with $\Delta\kappa_F \approx 11\%$ and $\Delta\kappa_V \approx 5\%$. It is interesting to note some historical context before concluding the combination results of the Higgs boson; initially, before the discovery of the Higgs boson, it was mentioned that "*at first look, the LHC is unlikely to ever get to 6% sensitivity* [102]." The LHC experiments have exceeded this expectation.

7.3 Combinations with other Electroweak Precision Observables

There is an enterprise of on-going work in constraining the SM parameter spaces through a global fits to several electroweak precision observables (EWPO) [103]. The measurements from several experiments from past decades are combined to constrain the SM parameter spaces. One might naively expect that the measurement of the Higgs properties through the direct production of the Higgs boson decays yields the

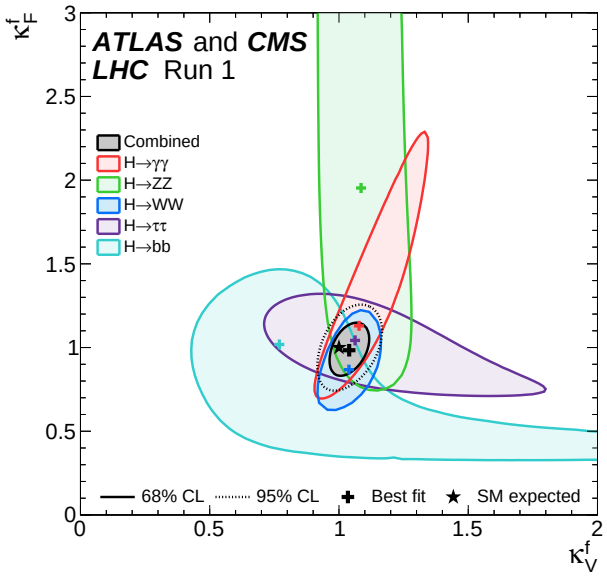


Figure 7.7: Constraint on κ_F vs. κ_V from the five major decay channels from ATLAS and CMS. The combined κ_F vs. κ_V constraint are shown along with the breakdown of each decay channel. The best constraint on both κ_F and κ_V is from $H \to WW^*$ [22].

most precise results. This is certainly true for the direct measurement, but through the usage of EWPO *indirect* measurement could shed light to the Higgs couplings with higher precision.

The coupling of the Higgs boson to the W and Z boson leads to higher order corrections to W and Z boson masses. Precise measurements of the vector boson masses can therefore constrain the Higgs boson couplings to the vector bosons. The relevant EWPOs are referred to as the S and T parameters named oblique parameters in the literature [104]

$$S = \frac{1}{12\pi}(1-\kappa_V^2)\ln\frac{\Lambda^2}{m_H^2}, \quad T = -\frac{3}{16\pi\cos\theta_{\text{eff}}^{\ell\ 2}}(1-\kappa_V^2)\ln\frac{\Lambda^2}{m_H^2}, \quad \Lambda = \frac{\lambda}{\sqrt{|1-\kappa_V^2|}}, \tag{7.8}$$

where the cut-off scale Λ represents the mass scale of the new states, and the $\theta_{\text{eff}}^{\ell}$ is the effective leptonic electroweak mixing angle. One should note that the higher the Λ, the smaller the κ_V. If the parameters S and T are away from 0, they modify the vector boson propagators. In the denominator of the logarithm is the mass of the Higgs boson, so the oblique parameters were used to predict the mass of the Higgs boson in the past.

Since the Higgs boson mass is now determined with high precision, the equation can be turned on its head to constrain κ_V. A private global fit is performed by the $Gfitter$ collaboration by combining existing EWPO measurements with the publicly available ATLAS and CMS results on the Higgs boson propeties [103]. The resulting constraint in the κ_F vs. κ_V plane are shown in Fig. 7.8 for $\lambda = 3$ TeV. The orange contours are

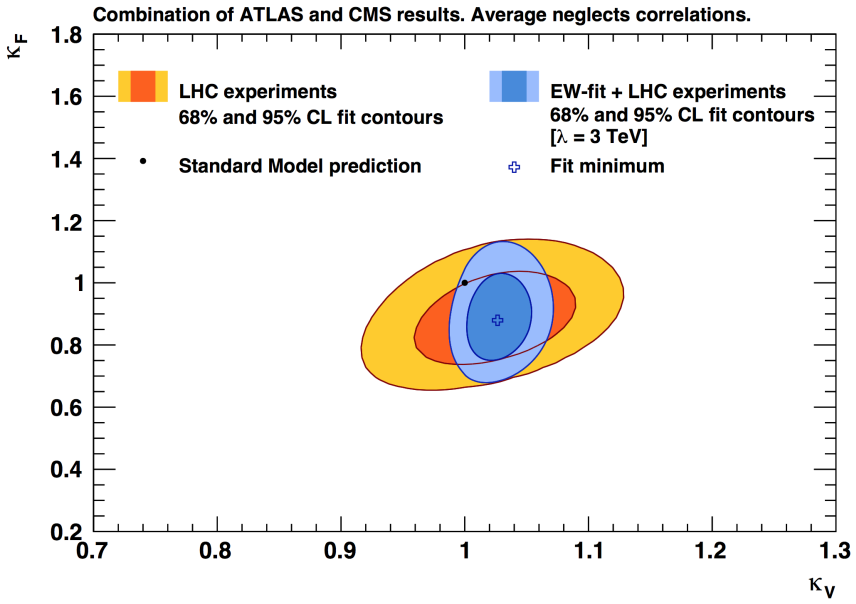


Figure 7.8: Constraint on κ_F vs. κ_V from the five major decay channels from ATLAS and CMS combined with all other existing electroweak precision observable (EWPO) measurements. The combined κ_F vs. κ_V constraint are shown for LHC experiment results only and LHC experiment plus other EWPO measurements. The κ_V sensitivity improves over the LHC only constraint by a significant factor [103].

the direct Higgs boson properties measurement and the blue contours are the combination of Higgs boson and EWPO measurements. The measured best-fit values of the κ_V for $\lambda = 1$ TeV, 3 TeV, 10 TeV are,

$$\kappa_V = 1.037^{+0.029}_{-0.026}, \quad 1.027^{+0.020}_{-0.019}, \quad 1.021^{+0.015}_{-0.014}, \tag{7.9}$$

showing a κ_V measurement improvement of factor of three or more compared to the direct Higgs boson properties measurement [103]. It has been shown that the majority of the sensitivity comes from the W boson mass measurement [105]; a precise indirect measurement from a copious amount of W boson events have won against the direct measurement of only a handful of Higgs boson events.

The κ_V may have been improved by the measurements on the EWPO but the κ_F does not improve. The EWPO does not have a strong dependence on κ_F, as the W and Z boson propagators receive corrections from κ_F at a much higher order with large suppressions. Thus the κ_F measurement must rely on the direct measurements of the Higgs boson couplings. This fact puts the importance of the VBF produced $H \to WW^*$ at the highest level among all decay channels as the measurement of VBF produced $H \to WW^*$ *is* the main driving force behind the sensitivity on κ_F; the *only* other channel of relevance is the ggF produced $H \to \tau\tau$, which is an experimentally more difficult final state. *All* other decay channel measurements only contribute to κ_V sensitivity and are irrelevant compared to the constraint from EWPO.

7.4 Suggestions for Future Work

In this section, few suggestions for future work to improve the analyses presented in Chapter 5 and Chapter 6 is documented.

7.4.1 VBF $H \to WW^* \to \ell\nu\ell\nu$

A public result of VBF $H \to WW^* \to \ell\nu\ell\nu$ analysis in Run 2 with the data corresponding to an integrated luminosity of 5.8 fb^{-1} at $\sqrt{s} = 13$ TeV recorded by the ATLAS detector has been made public [107]. The analysis strategy is almost identical to the one presented in Chapter 5. The Table 5.7 shows that, among the major sources of uncertainties, (1) "Signal VBF acceptance", (2) "Background WW", (3) "Background top-quark", and (4) "Jets" to be significantly large. This section will briefly discuss each systematics.

The signal VBF acceptance uncertainties are dominated by the difference observed in HERWIG and PYTHIA modeling of the underlying event. As it was discussed in Sec. 5.5, the usage of $p_{\mathrm{T}}^{\mathrm{sum}}$ variable in the BDT algorithm causes the BDT shape to be sensitive to the soft jet activities. A large uncertainty induced by the inclusion of $p_{\mathrm{T}}^{\mathrm{sum}}$ can be easily avoided by not including the $p_{\mathrm{T}}^{\mathrm{sum}}$ variable to the BDT algorithm. It has been shown that the $p_{\mathrm{T}}^{\mathrm{sum}}$ brings about $\sim 5\%$ improvement in significance, so the suggestion is to reassess when a relatively large amount of data is taken.

The uncertainties in "Background WW" are theoretical ones. As discussed in Sec. 5.4.2, the difficulty of establishing a phase-space region with relatively high purity of WW prohibited from setting up a control region. But as it was shown with the usage of m_{T2} variable, a relatively pure phase-space region can be selected. The suggestion is to use the region with high m_{T2} as a control region. This will exchange the theoretical uncertainty with a statistical uncertainty, which is always understood better.

The uncertainties in "Background top-quark" are also theoretical ones. Although the fractional theoretical uncertainties measured for the top-quark background is around only $10 - 20\%$, relatively large contribution of the background in the fit region affects the measurement in a significant way. The top-quark process cross-section is expected to rise from $sqrts = 8$ TeV collisions to 13 TeV collisions more rapidly than most other SM processes. A more sophisticated background suppression method will be useful in reducing the uncertainties on the μ_{VBF} measurement.

A truth level study is conducted to understand what types of top-quark background process enters in to the high O_{BDT} score regions. For most events entering the signal regions, the one of the b-jets are found to be outside of the tracking region, failing the b-tagging identification. To further reduce the kind of events with b-jets outside of the tracking region, an inner detector upgrade covering high η region may be a solution. For events that have b-jets inside the tracking regions, a large fraction of events are removed by the b-tagging

identification, and only a few fraction of the events are left. These types of events can be further reduced with better b-tagging algorithms. There is however a third type of events, which are worth investigating. The third category of events are where one of the b-quark originating jets are simply not reconstructed. These types of events are found to be around a quarter of all the events entering the signal region. More interestingly, more than half of these b-quarks that are not reconstructed do in fact carry a $p_{\mathrm{T}} > 20$ GeV. Increasing the reconstruction efficiency on these b-quarks can allow the analyzers to reject more top-quark backgrounds by inferring the presence of b-jets.

Lastly, the uncertainty on the "Jet" category comes from the difference between HERWIG and PYTHIA in extrapolating the calibration scale factors from central regions to the forward regions [45]. It may be worthwhile developing a dedicated analysis measuring the calibration scale factors for the forward jets directly with higher precision.

7.4.2 $H \to ss \to bb\mu\mu$

The major source of inefficiency in signal acceptance comes from the jet reconstruction. For the signal sample with $m_s = 60$ GeV, once the requirement of $n_b \geq 2$ are applied, 82% of events are lost. The fraction of events that satisfy $n_b = 1$ is 48% and the fraction of events that satisfy $n_b = 0$ is 34%. The main reason for the loss of efficiency is related to the fact that the subleading b-jets are very soft. Fig. 7.9 shows the truth level subleading b-quark p_{T} distribution from a signal sample with $m_s = 20$ GeV. A dashed line is drawn at 20 GeV to indicate the the jet p_{T} threshold applied to the reconstructed jets. Only the b-quarks from the tail of the distribution are likely to be reconstructed.

The lowering of the jet p_{T} thresholds without introducing a large number of pile-up jets is a difficult task. As a solution, increasingly other Run 2 analyses are using jets reconstructed solely from tracks, referred to as track-jets. With the pinpointing capability of the tracks, track-jets can be associated to primary vertices of origin with high precision. This in return allows for a lowered thresholds on the track-jet p_{T}. In ATLAS, a p_{T} thresholds as low as 7 GeV has been explored. It is important to note that the track-jets tend to be off by factor of two in p_{T} due to missed neutral components. The requirement of $p_{\mathrm{T}} > 7$ GeV can be interpreted as $p_{\mathrm{T}} > 15$ GeV in terms of calorimeter-based jet. This effective 5 GeV gain in lowered p_{T} thresholds helps to recover lost signal events. Of the 48% of events that only has one b-jets in the events, a fraction of 13% of events have at least one track-jet with $p_{\mathrm{T}} > 7$ GeV. Compared to the existing 18% signal acceptance, this adds about 50% more signal events to the signal region.

The recovered events through reconstructing one calorimeter-based jet (calo-jets) plus a one track-jet, exhibits good mass resonance structure. Fig. 7.10 shows the invariant mass of the four-fermion system for

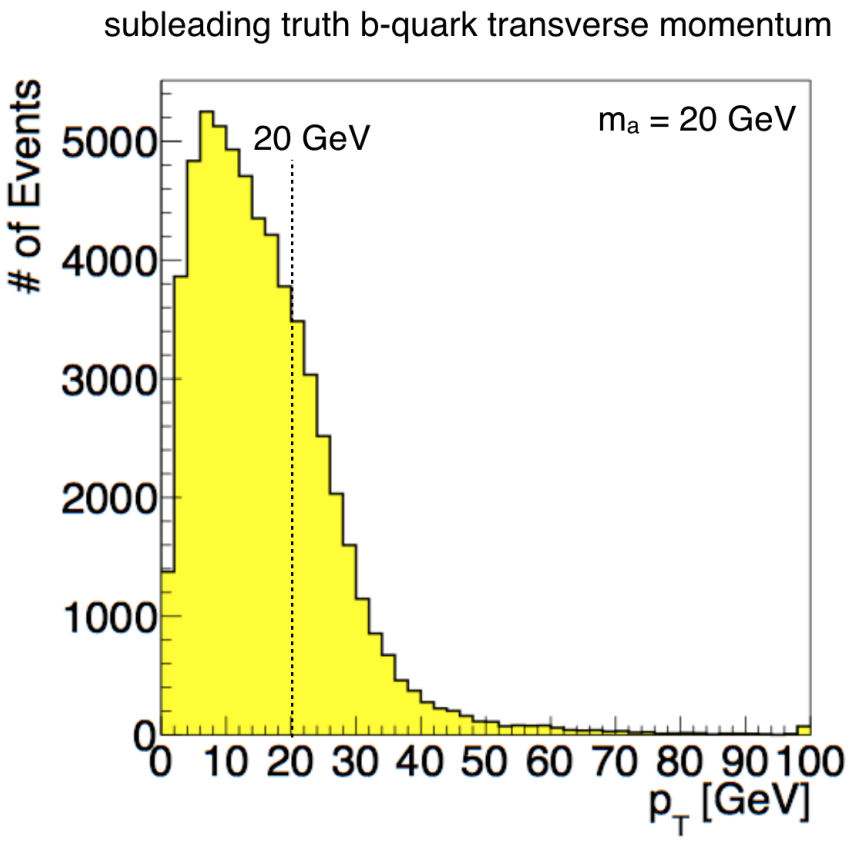


Figure 7.9: The truth level subleading b-quark p_{T} distribution from a signal sample with $m_s = 20$ GeV. A dashed line is drawn to indicate the jet p_{T} threshold applied to the reconstructed jets.

five signal samples. For all mass points, the reconstructed Higgs boson resonance peak exhibits a sharp structure. The mean value, however, are slightly shifted towards a lower value. This can be understood as due to the missing neutral components in track-jet reconstruction.

By studying the track-jet reconstruction, one can formulate a correction function to shift back the loss portions of the neutral contributions. Such work is on-going within ATLAS collaboration.

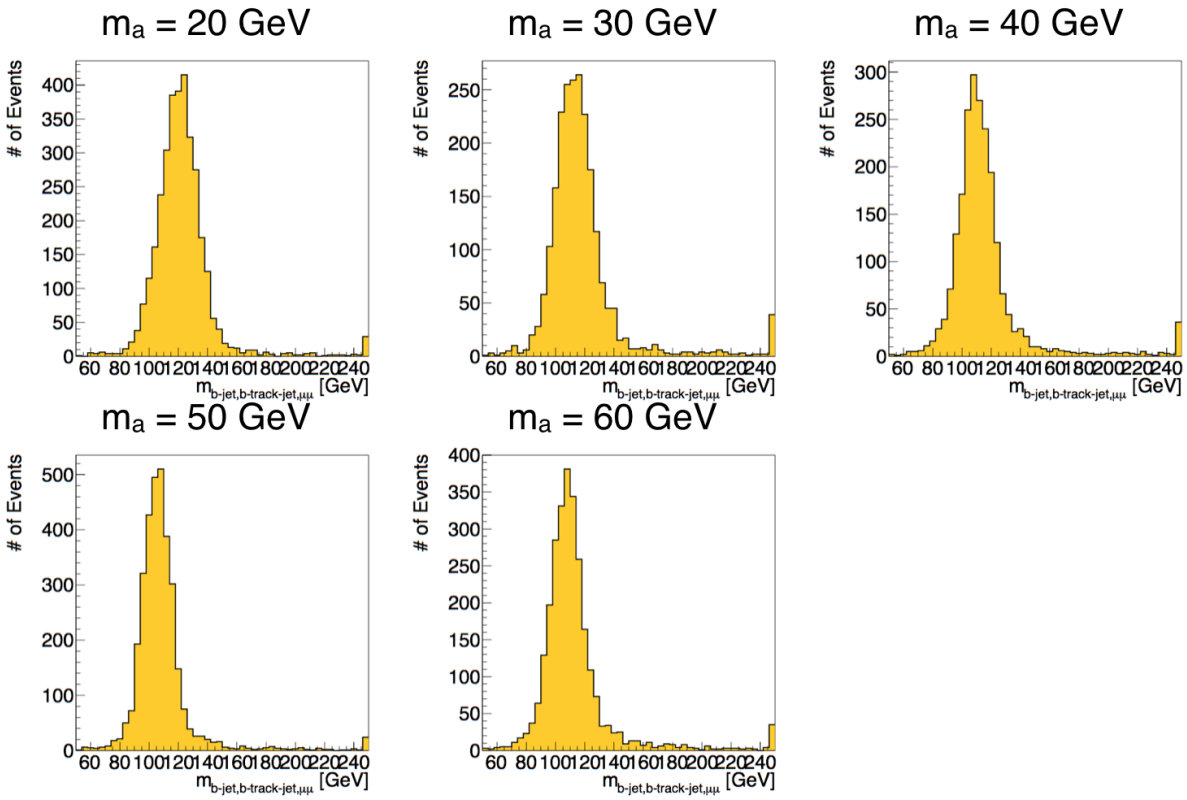


Figure 7.10: Four-fermion mass distribution formed by two reconstructed muons plus one track-jet and one calo-jet. For all signal mass points, the events show a sharp peak structure, but with a shifted mean.

Chapter 8

Conclusions

This thesis presented the first evidence for the vector boson fusion (VBF) produced Higgs boson in exclusive Higgs boson decay channel ($H \to WW^*$). The analysis is performed using an integrated luminosity of 25 fb^{-1} from $\sqrt{s} = 7$ and 8 TeV pp collision data recorded by the ATLAS detector at the LHC. The subsequent measurement of the Higgs couplings in this channel showed best constraint on the fermionic coupling of the Higgs boson.

With the existence of the Higgs sector established, the Higgs boson is used as a tool for discovery. Many models predict possible connection of Higgs boson to DM. In particular, models with an additional light pseudoscalar boson have been proposed to explain the DM. If true, the Higgs boson may shed light on the DM through a non-SM decay of the Higgs boson.

The thesis also presented the strategy of the search for a non-SM decay of the Higgs boson $H \to ss \to bb\mu\mu$, and the analysis of the expected sensitivity. The expected upper limit on the branching ratio is found to be $\mathcal{B}(H \to bb\mu\mu) \leq 2.4-4.8\times10^{-4}$ at 95% confidence limit across the mass range of 20 GeV $< m_a <$ 60 GeV. The expected limit corresponds to tighter limit than the indirect constraint on $\mathcal{B}(H \to BSM)$ from Higgs coupling measurements.

References

[1] Q. R. Ahmad *et al.* [SNO Collaboration], Phys. Rev. Lett. **89**, 011301 (2002) doi:10.1103/PhysRevLett.89.011301 [nucl-ex/0204008].

[2] The highest point shown around October 15th shows data point at 51. https://atlas.web.cern.ch/Atlas/GROUPS/DATAPREPARATION/PublicPlots/2016/DataSummary/figs/peakMuByFill.pdf

[3] A. D. Martin, Acta Phys. Polon. B **39**, 2025 (2008) [arXiv:0802.0161 [hep-ph]].

[4] C. Patrignani *et al.* [Particle Data Group], Chin. Phys. C **40**, no. 10, 100001 (2016). doi:10.1088/1674-1137/40/10/100001

[5] ATLAS Collaboration, *Summary plots from the ATLAS Standard Model physics group*.

[6] F. Englert and R. Brout, Phys. Rev. Lett. **13**, 321 (1964). doi:10.1103/PhysRevLett.13.321

[7] P. W. Higgs, Phys. Lett. **12**, 132 (1964). doi:10.1016/0031-9163(64)91136-9

[8] P. W. Higgs, Phys. Rev. Lett. **13**, 508 (1964). doi:10.1103/PhysRevLett.13.508

[9] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, Phys. Rev. Lett. **13**, 585 (1964). doi:10.1103/PhysRevLett.13.585

[10] P. W. Higgs, Phys. Rev. **145**, 1156 (1966). doi:10.1103/PhysRev.145.1156

[11] T. W. B. Kibble, Phys. Rev. **155**, 1554 (1967). doi:10.1103/PhysRev.155.1554

[12] J. R. Ellis, M. K. Gaillard and D. V. Nanopoulos, Nucl. Phys. B **106**, 292 (1976). doi:10.1016/0550-3213(76)90382-5

[13] S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967). doi:10.1103/PhysRevLett.19.1264

[14] S. L. Glashow, Nucl. Phys. **22**, 579 (1961). doi:10.1016/0029-5582(61)90469-2

[15] A. Salam, in *Proceedings of the Nobel Symposium*, edited by N. Svartholm (Almqvist & Wiksell, Stockholm, 1968), p. 367.

[16] M. S. Neubauer, Ann. Rev. Nucl. Part. Sci. **61**, 223 (2011). doi:10.1146/annurev-nucl-102010-130106

[17] J. Bagger, V. D. Barger, K. m. Cheung, J. F. Gunion, T. Han, G. A. Ladinsky, R. Rosenfeld and C. P. Yuan, Phys. Rev. D **49**, 1246 (1994) doi:10.1103/PhysRevD.49.1246 [hep-ph/9306256].

[18] G. Aad *et al.* [ATLAS Collaboration], Phys. Rev. D **92**, no. 1, 012006 (2015) doi:10.1103/PhysRevD.92.012006 [arXiv:1412.2641 [hep-ex]].

[19] G. Aad *et al.* [ATLAS Collaboration], Phys. Rev. Lett. **113**, no. 14, 141803 (2014) doi:10.1103/PhysRevLett.113.141803 [arXiv:1405.6241 [hep-ex]].

[20] Y. H. Qi, Y. P. Kuang, B. J. Liu and B. Zhang, Phys. Rev. D **79**, 055010 (2009) Erratum: [Phys. Rev. D **82**, 119902 (2010)] doi:10.1103/PhysRevD.82.119902 [arXiv:0811.3099 [hep-ph]].

[21] K. Cheung, C. W. Chiang and T. C. Yuan, Phys. Rev. D **78**, 051701 (2008) doi:10.1103/PhysRevD.78.051701 [arXiv:0803.2661 [hep-ph]].

[22] G. Aad *et al.* [ATLAS and CMS Collaborations], JHEP **1608**, 045 (2016) doi:10.1007/JHEP08(2016)045 [arXiv:1606.02266 [hep-ex]].

[23] J. Ellis and T. You, JHEP **1209**, 123 (2012) doi:10.1007/JHEP09(2012)123 [arXiv:1207.1693 [hep-ph]].

[24] G. Aad *et al.* [ATLAS Collaboration], JHEP **1511**, 206 (2015) doi:10.1007/JHEP11(2015)206 [arXiv:1509.00672 [hep-ex]].

[25] S. Dittmaier *et al.* [LHC Higgs Cross Section Working Group], doi:10.5170/CERN-2011-002 arXiv:1101.0593 [hep-ph].

[26] D. Stojkovic, Int. J. Mod. Phys. D **22**, 1342017 (2013) doi:10.1142/S0218271813420170 [arXiv:1305.6960 [gr-qc]].

[27] C. Csaki, Mod. Phys. Lett. A **11**, 599 (1996) doi:10.1142/S021773239600062X [hep-ph/9606414].

[28] B. Patt and F. Wilczek, hep-ph/0605188.

[29] D. Curtin and C. B. Verhaaren, JHEP **1512**, 072 (2015) doi:10.1007/JHEP12(2015)072 [arXiv:1506.06141 [hep-ph]].

[30] D. Curtin *et al.*, Phys. Rev. D **90**, no. 7, 075004 (2014) doi:10.1103/PhysRevD.90.075004 [arXiv:1312.4992 [hep-ph]].

[31] G. Aad *et al.* [ATLAS Collaboration], JINST **3**, S08003 (2008). doi:10.1088/1748-0221/3/08/S08003

[32] CISCO, *Cisco Visual Networking Index: Forecast and Methodology, 20152020*

[33] P. O. Hansson Adrian, arXiv:1111.4190 [hep-ex].

[34] Y. Nakahama, J. Phys. Conf. Ser. **664**, no. 8, 082037 (2015). doi:10.1088/1742-6596/664/8/082037

[35] M. Shochet *et al.* [FTK group], CERN-LHCC-2013-007. ATLAS-TDR-021 (2013).

[36] J. Alison *et al.* [FTK group], ATL-DAQ-INT-2014-005, Geneva, 2014

[37] G. Aad *et al.* [ATLAS Collaboration], Eur. Phys. J. C **74**, no. 11, 3130 (2014) doi:10.1140/epjc/s10052-014-3130-x [arXiv:1407.3935 [hep-ex]].

[38] G. Aad *et al.* [ATLAS Collaboration], Eur. Phys. J. C **72**, 1909 (2012) doi:10.1140/epjc/s10052-012-1909-1 [arXiv:1110.3174 [hep-ex]].

[39] [ATLAS Collaboration], ATLAS-CONF-2012-047.

[40] The ATLAS collaboration [ATLAS Collaboration], ATLAS-CONF-2014-032.

[41] M. Cacciari and G. P. Salam, Phys. Lett. B **641**, 57 (2006) doi:10.1016/j.physletb.2006.08.037 [hep-ph/0512210].

[42] M. Cacciari, G. P. Salam and G. Soyez, JHEP **0804**, 063 (2008) doi:10.1088/1126-6708/2008/04/063 [arXiv:0802.1189 [hep-ph]].

[43] W. Lampl *et al.*, ATL-LARG-PUB-2008-002, ATL-COM-LARG-2008-003.

[44] G. Aad *et al.* [ATLAS Collaboration], Eur. Phys. J. C **73**, no. 3, 2304 (2013) doi:10.1140/epjc/s10052-013-2304-2 [arXiv:1112.6426 [hep-ex]].

[45] G. Aad *et al.* [ATLAS Collaboration], Eur. Phys. J. C **75**, 17 (2015) doi:10.1140/epjc/s10052-014-3190-y [arXiv:1406.0076 [hep-ex]].

[46] [ATLAS Collaboration], ATLAS-CONF-2012-064.

[47] The ATLAS collaboration [ATLAS Collaboration], ATLAS-CONF-2013-083.

[48] The ATLAS collaboration [ATLAS Collaboration], ATLAS-CONF-2014-046.

[49] [ATLAS Collaboration], ATLAS-CONF-2011-102.

[50] The ATLAS collaboration [ATLAS Collaboration], ATLAS-CONF-2014-004.

[51] G. Aad *et al.* [ATLAS Collaboration], Eur. Phys. J. C **72**, 1844 (2012) doi:10.1140/epjc/s10052-011-1844-6 [arXiv:1108.5602 [hep-ex]].

[52] ATLAS Collaboration [ATLAS Collaboration], ATL-PHYS-PUB-2010-013, ATL-COM-PHYS-2010-838.

[53] S. Agostinelli *et al.* [GEANT4 Collaboration], Nucl. Instrum. Meth. A **506**, 250 (2003). doi:10.1016/S0168-9002(03)01368-8

[54] P. Nason, JHEP **0411**, 040 (2004) doi:10.1088/1126-6708/2004/11/040 [hep-ph/0409146].

[55] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau and A. D. Polosa, JHEP **0307**, 001 (2003) doi:10.1088/1126-6708/2003/07/001 [hep-ph/0206293].

[56] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert and J. Winter, JHEP **0902**, 007 (2009) doi:10.1088/1126-6708/2009/02/007 [arXiv:0811.4622 [hep-ph]].

[57] B. P. Kersevan and E. Richter-Was, Comput. Phys. Commun. **184**, 919 (2013) doi:10.1016/j.cpc.2012.10.032 [hep-ph/0405247].

[58] N. Kauer, JHEP **1312**, 082 (2013) doi:10.1007/JHEP12(2013)082 [arXiv:1310.7011 [hep-ph]].

[59] T. Sjostrand, S. Mrenna and P. Z. Skands, JHEP **0605**, 026 (2006) doi:10.1088/1126-6708/2006/05/026 [hep-ph/0603175].

[60] T. Sjostrand, S. Mrenna and P. Z. Skands, Comput. Phys. Commun. **178**, 852 (2008) doi:10.1016/j.cpc.2008.01.036 [arXiv:0710.3820 [hep-ph]].

[61] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour and B. R. Webber, JHEP **0101**, 010 (2001) doi:10.1088/1126-6708/2001/01/010 [hep-ph/0011363].

[62] J. M. Butterworth, J. R. Forshaw and M. H. Seymour, Z. Phys. C **72**, 637 (1996) doi:10.1007/BF02909195, 10.1007/s002880050286 [hep-ph/9601371].

[63] H. L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin and C.-P. Yuan, Phys. Rev. D **82**, 074024 (2010) doi:10.1103/PhysRevD.82.074024 [arXiv:1007.2241 [hep-ph]].

[64] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. M. Nadolsky and W. K. Tung, JHEP **0207**, 012 (2002) doi:10.1088/1126-6708/2002/07/012 [hep-ph/0201195].

[65] [ATLAS Collaboration], ATL-PHYS-PUB-2012-003, ATL-COM-PHYS-2012-738.

[66] A. D. Martin, W. J. Stirling, R. S. Thorne and G. Watt, Eur. Phys. J. C **63**, 189 (2009) doi:10.1140/epjc/s10052-009-1072-5 [arXiv:0901.0002 [hep-ph]].

[67] L. Breiman, J. Friedman, R. Olshen, and C. Stone, Classification and Regression Trees (Chapman and Hall, New York, 1984); Y. Freund and R. E. Schapire, J. Com- put. Syst. Sci. 55, 119 (1997); J. Friedman, Computational Statistics and Data Analysis 38, 267 (2002).

[68] A. Hocker *et al.*, PoS ACAT , 040 (2007) [physics/0703039 [PHYSICS]].

[69] C. A. Nelson, Phys. Rev. D **37**, 1220 (1988). doi:10.1103/PhysRevD.37.1220

[70] M. Dittmar and H. K. Dreiner, Phys. Rev. D **55**, 167 (1997) doi:10.1103/PhysRevD.55.167 [hep-ph/9608317].

[71] A. J. Barr, B. Gripaios and C. G. Lester, JHEP **0907**, 072 (2009) doi:10.1088/1126-6708/2009/07/072 [arXiv:0902.4864 [hep-ph]].

[72] R. K. Ellis, I. Hinchliffe, M. Soldate and J. J. van der Bij, Nucl. Phys. B **297**, 221 (1988). doi:10.1016/0550-3213(88)90019-3

[73] T. Plehn, D. L. Rainwater and D. Zeppenfeld, Phys. Rev. D **61**, 093005 (2000) doi:10.1103/PhysRevD.61.093005 [hep-ph/9911385].

[74] G. Aad *et al.* [ATLAS Collaboration], arXiv:0901.0512 [hep-ex], p. 1280. ATLAS Collaboration, arXiv:0901.0512, p. 1280.

[75] V. D. Barger, R. J. N. Phillips and D. Zeppenfeld, Phys. Lett. B **346**, 106 (1995) doi:10.1016/0370-2693(95)00008-9 [hep-ph/9412276].

[76] B. Cerio, CERN-THESIS-2015-232.

[77] H. C. Cheng and Z. Han, JHEP **0812**, 063 (2008) doi:10.1088/1126-6708/2008/12/063 [arXiv:0810.5178 [hep-ph]].

[78] I. W. Stewart and F. J. Tackmann, Phys. Rev. D **85**, 034011 (2012) doi:10.1103/PhysRevD.85.034011 [arXiv:1107.2117 [hep-ph]].

[79] HWW Working Group, *Background esimates in the* $H \to WW^* \to \ell\nu\ell\nu$ *analysis with* 20 fb^{-1} *of data collected with the ATLAS detector at* $\sqrt{s} = 8$ TeV *— Supporting Note,* ATL-COM-PHYS-2013-1630, CERN, Geneva, April 2014.

[80] HWW Working Group, *Analysis of* $H \to WW^* \to \ell\nu\ell\nu$ *ggF and VBF production modes with* 20 fb^{-1} *and* 5 fb^{-1} *of data collected with the ATLAS detector at* $\sqrt{s} = 8$ *and* 7 TeV *— Supporting Note,* ATL-COM-PHYS-2014-466 CERN, Geneva, May 2014.

[81] G. Aad *et al.* [ATLAS Collaboration], Phys. Rev. D **90**, no. 5, 052004 (2014) doi:10.1103/PhysRevD.90.052004 [arXiv:1406.3827 [hep-ex]].

[82] LHC Higgs Cross Section Working Group, *In preparation*, https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageAt13TeV

[83] S. Dawson *et al.*, arXiv:1310.8361 [hep-ex].

[84] S. Ipek, D. McKeen and A. E. Nelson, Phys. Rev. D **90**, no. 5, 055021 (2014) doi:10.1103/PhysRevD.90.055021 [arXiv:1404.3716 [hep-ph]].

[85] D. Curtin, R. Essig and Y. M. Zhong, JHEP **1506**, 025 (2015) doi:10.1007/JHEP06(2015)025 [arXiv:1412.4779 [hep-ph]].

[86] C. Anastasiou, C. Duhr, F. Dulat, F. Herzog and B. Mistlberger, Phys. Rev. Lett. **114**, 212001 (2015) doi:10.1103/PhysRevLett.114.212001 [arXiv:1503.06056 [hep-ph]].

[87] C. Anastasiou, C. Duhr, F. Dulat, E. Furlan, T. Gehrmann, F. Herzog, A. Lazopoulos and B. Mistlberger, JHEP **1605**, 058 (2016) doi:10.1007/JHEP05(2016)058 [arXiv:1602.00695 [hep-ph]].

[88] S. Actis, G. Passarino, C. Sturm and S. Uccirati, Phys. Lett. B **670**, 12 (2008) doi:10.1016/j.physletb.2008.10.018 [arXiv:0809.1301 [hep-ph]].

[89] C. Anastasiou, R. Boughezal and F. Petriello, JHEP **0904**, 003 (2009) doi:10.1088/1126-6708/2009/04/003 [arXiv:0811.3458 [hep-ph]].

[90] J. Butterworth *et al.*, J. Phys. G **43**, 023001 (2016) doi:10.1088/0954-3899/43/2/023001 [arXiv:1510.03865 [hep-ph]].

[91] G. Aad *et al.* [ATLAS Collaboration], Eur. Phys. J. C **70**, 823 (2010) doi:10.1140/epjc/s10052-010-1429-9 [arXiv:1005.4568 [physics.ins-det]].

[92] J. Erdmann, S. Guindon, K. Kroeninger, B. Lemmer, O. Nackenhorst, A. Quadt and P. Stolte, Nucl. Instrum. Meth. A **748**, 18 (2014) doi:10.1016/j.nima.2014.02.029 [arXiv:1312.5595 [hep-ex]].

[93] P. Chang, C. Hayes, J. Hobbs, B. Kaplan, A. Liu, L. Morvaj, M. Neubauer, Y. Zhang *Search for a Higgs boson decaying to pseudoscalar pairs in the bbμμ final state at* $\sqrt{s} = 13$ TeV *with ATLAS detector. — Supporting Note,* ATL-COM-PHYS-2015-1553 CERN, Geneva, Dec 2015.

[94] M. Baak, G. J. Besjes, D. Cte, A. Koutsman, J. Lorenz and D. Short, Eur. Phys. J. C **75**, 153 (2015) doi:10.1140/epjc/s10052-015-3327-7 [arXiv:1410.1280 [hep-ex]].

[95] T. Junk, Nucl. Instrum. Meth. A **434**, 435 (1999) doi:10.1016/S0168-9002(99)00498-2 [hep-ex/9902006].

[96] A. L. Read, J. Phys. G **28**, 2693 (2002). doi:10.1088/0954-3899/28/10/313

[97] G. Cowan, K. Cranmer, E. Gross and O. Vitells, Eur. Phys. J. C **71**, 1554 (2011) Erratum: [Eur. Phys. J. C **73**, 2501 (2013)] doi:10.1140/epjc/s10052-011-1554-0, 10.1140/epjc/s10052-013-2501-z [arXiv:1007.1727 [physics.data-an]].

[98] G. Aad *et al.* [ATLAS Collaboration], Eur. Phys. J. C **76**, no. 1, 6 (2016) doi:10.1140/epjc/s10052-015-3769-y [arXiv:1507.04548 [hep-ex]].

[99] ATLAS Collaboration, *Summary plots from the ATLAS Higgs physics group.*

[100] V. Khachatryan *et al.* [CMS Collaboration], Eur. Phys. J. C **75**, no. 5, 212 (2015) doi:10.1140/epjc/s10052-015-3351-7 [arXiv:1412.8662 [hep-ex]].

[101] S. Chatrchyan *et al.* [CMS Collaboration], JHEP **1401**, 096 (2014) doi:10.1007/JHEP01(2014)096 [arXiv:1312.1129 [hep-ex]].

[102] R. S. Gupta, H. Rzehak and J. D. Wells, Phys. Rev. D **86**, 095001 (2012) doi:10.1103/PhysRevD.86.095001 [arXiv:1206.3560 [hep-ph]].

[103] M. Baak *et al.* [Gfitter Group], Eur. Phys. J. C **74**, 3046 (2014) doi:10.1140/epjc/s10052-014-3046-5 [arXiv:1407.3792 [hep-ph]].

[104] M. E. Peskin and T. Takeuchi, Phys. Rev. D **46**, 381 (1992). doi:10.1103/PhysRevD.46.381

[105] M. Ciuchini, E. Franco, S. Mishima and L. Silvestrini, JHEP **1308**, 106 (2013) doi:10.1007/JHEP08(2013)106 [arXiv:1306.4644 [hep-ph]].

[106] J. R. Espinosa, C. Grojean, M. Muhlleitner and M. Trott, JHEP **1212**, 045 (2012) doi:10.1007/JHEP12(2012)045 [arXiv:1207.1717 [hep-ph]].

[107] The ATLAS collaboration [ATLAS Collaboration], ATLAS-CONF-2016-112.

[108] T. Han, G. Valencia and S. Willenbrock, Phys. Rev. Lett. **69**, 3274 (1992) doi:10.1103/PhysRevLett.69.3274 [hep-ph/9206246].